

Unit
04

INFORMATION HANDLING

Q.1 Define the following terms.

(i) Data

The numerical figures obtained from any field of study are known as data.

(ii) Primary data

The data directly collected from its source is called primary data.

(iii) Secondary data

The data which have been passed through some statistical treatments at least one is called secondary data.

(iv) Arithmetic Mean

It is defined as a value which is obtained by dividing the sum of all the values by their number of observations, it is denoted by \bar{x} .

$$\bar{x} = \frac{\sum X}{n}$$

(v) Merits of Arithmetic Mean.

- (i) It is most simple form of an average.
- (ii) It is easy to find and understand.
- (iii) To find this, we need only the sum of all the observations and number of observation

(vi) Demerits of Arithmetic Mean

- (i) It is highly affected by extreme values.
- (ii) It provides a high value on having one very large value in the data set.

(vii) Merits of Median

- i) It is easy to find and interpret.
- i) It is not affected by extreme values.

(viii) Demerits of median

- i) It cannot be calculated unless the data are ranged in some order.
-) It is not based on all items.

(ix) Merits of mode

- (i) In certain cases mode can be located easily in the data set.
- (ii) It is most common value.

(x) Demerits of mode.

- (i) Sometimes, data do not contain any mode.
- (ii) Data set may contain more than one mode in this situation mode is not most suitable measure.
- (iii) Mode cannot be easily calculated from the grouped data unless data are very large.

(xi) Class limits

Each group or class is defined by two values, one small and other large the smaller one is known as lower limit and larger one is known as upper limit. e.g. in (60-62), the lower limit is 60 and the upper limit is 62.

(xii) Class boundaries

Class boundaries are the actual lower and upper class limits is a class.

Example:

In Class interval 40 – 45, class boundaries are 39.5 – 45.5 smaller number 39.5 is known as lower class boundary where as larger number 45.5 is known as upper class boundary.

(xiii) Class frequency

Ans: The number of occurrences of items corresponding to the class interval is known as class frequency.

(xiv) Frequency distribution

A method of organizing and condensing the raw data is the Conversion of raw data into a grouped data. The various items of data are classified into certain groups or classes and numbers of items lying in each group or class

is put against that group or class. The data organized and summarized in this way is known as frequency distribution.

(xv) Measures of Dispersion

The measure which measure the spreadness of values from central value (average) are called measured of dispersion.

(xvi) Co-efficient of Variation. (C.V)

To compare the variation of two or more series measured in different unit a measure of relative dispersion, co-efficient of variation is used. It is expressed in terms of percentage given by ratio of standard deviation S.D and its mean (\bar{X})

$$C.V = \frac{S}{\bar{X}} \times 100$$

Note: Smaller Co-efficient of variation is more consistent in performance.

(xvii) Types of dispersion

It has the following types.

- (i) range (ii) Variance
- (iii) Standard deviation (iv) Mean deviation
- (v) Quartile deviation.

(xviii) Range

It is obtained by taking the difference between the largest and the smallest value in a given data. It is mathematically defined as

$$\text{Range (R)} = X_m - X_1$$

Where X_m = largest value

X_1 = smallest value

(xix) Variance

It is a single value, obtained by dividing the sum of squares of deviations taken from \bar{X} by the number of observations in data set. It is denoted by s^2 .

$$s^2 = \frac{\sum (x - \bar{x})^2}{n} \text{ or } s^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

(xx) Standard deviation

The standard deviation is the positive square root of variance it is denoted by s .

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

$$\text{or } s = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2}$$

(xxi) Constant:

Any fixed quantity that has a single value is known as a constant.

e.g. any real number can be taken as constant.

(xxii) Variable:

Any characteristic whose values are always different from one individual to another is known as a variable. e.g. weights of students

(xxiii) Discrete Variable:

It can take only some specific values present in the data. e.g. The number of children in a family.

(xxiv) Continuous Variable:

It can take every possible value in a given interval say (a to b). It may be a whole figure or a fraction. e.g. Product of milk by cow.

(xxv) Difference between discrete and continuous variable

- (a) The number of students in a class (Discrete variable)
- (b) Milk produced by a cow. (Continuous variable)
- (c) The number of heads in tossing two coins (Discrete variable)
- (d) The number of sixes in rolling 4 dice (Discrete variable)
- (e) The age of an employee of an office (Continuous variable)
- (f) The life time of T.V tubes (Continuous Variable)
- (g) The number of prize bonds sold every day (Discrete variable)

(xxvi) Ungrouped data: Numerical facts which are obtained on the first hand and

recorded as they stand are known as ungrouped data.

(xxvii) Grouped data:

When the data have gone through some statistical process i.e. data may be classified into certain groups or into rows and columns.

(xxviii) Types of data

There are two main sources of data.

- i) Primary data
- ii) Secondary data

(xxix) Information Handling

To know about something is known as "information" and to present that information in a manageable way so that useful conclusions can be drawn is called information handling.

(xxx) Classification

Classification is a process of arranging the data into certain groups or classes having similar characteristics.

(xxxi) Class interval

The size, width or length of class interval is the difference between the upper (or lower) limits of any two consecutive classes. It is denoted by h .

(xxxii) Class mark

It is mid point of any class. It is the average value of the lower and upper class limits. E.g. in (60-62) the midpoint of group =

$$(x) = \frac{60+62}{2} = 61$$

(xxxiii) Tabulation

Tabulation of the data means to present the data in a classified form or into rows and columns. In this way the large number of observation is presented in a compact form which can be easily handled and analyzed.

(xxxiv) One-way Tabulation:

In one way tabulation, we present information only for one characteristic or variable.

(xxxv) Two-way Tabulation:

When we classify the data by two characteristics at the same time, it is known as two-way tabulation.

(xxxvi) Measure of central tendency

Techniques of univariate quantitative data analysis which is known as to be study of central tendency of the data. Study of central tendency means try to find single numerical value which may be used to represent the whole data or distribution. The central tendency of any data also known as the average value or measure of location.

(xxxvii) Types of measure of central tendency

Measure of central tendency commonly known as an average, has following types.

- i) Arithmetic mean
- ii) Median
- iii) Mode
- iv) Geometric mean
- v) Harmonic mean

(xxxviii) Median

When the data are arranged in some order. Median is the value which divides the data into two parts i.e. 50% of the data are before the median value and 50% after it. OR The value which divides the whole arranged data into two equal parts is called Median.

(xxxix) Mode

In a set of data, it is a value of the variable which appears or occurs most often. It is the most common value.

Q.2 How many classes or groups should be made to condense the raw data at hand without losing the important information contained there in.

Ans: There is no hard and fast rule for finding exact number of classes. However, classes or groups should be between 5 and 15. Too small number of groups results into loss of important information.

Q.3 Who collects data? What must be qualities that person who collects data?

Ans: A person who collects the data is called investigator must have the following quantities.

- (i) Intelligent, reliable and responsible.
- (ii) Properly trained and Polite
- (iii) Experienced, tactful and well aware of with the object he is dealing with

Q.4 Write types of diagram that can be used to represent a frequency distribution on a graph.

Ans: The following are the types of diagram that can be used to represent a frequency distribution on a graph.

- i) Histogram
- ii) Frequency polygon

Q.5 How do we calculate median for grouped data.

Ans: Median is obtained by the following formula.

$$\text{Median} = \ell + \frac{h}{f} \left(\frac{n}{2} - c \right)$$

Where

ℓ = Lower class boundary of median class

h = Length of class interval of median class

f = frequency of the median class

n = Total frequency i.e. $\sum f$.

c = Cumulative frequency preceding the median class.

Q.6 How do we calculate mode for grouped data.

Ans: Mode can be calculated by the following formula.

$$\text{Mode} = \ell + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

Where ℓ = Lower class boundary of the modal class

f_m = Frequency of the modal class

f_1 = Frequency preceding the modal class

f_2 = Frequency following the modal class

h = Class interval.

Q.7 Write situation when mean, median and mode are used.

Ans: Mean is used for general purpose i.e. to find average, such as weights, heights, distribution of grades etc.

Median is used for distribution of monthly or yearly incomes, wages etc.

Mode is used when most common value is required such as size of collars, size of shoes etc.

Q.8 What is an average ? Explain the method of calculation of arithmetic mean for ungrouped and as well as for grouped data.

Ans: **Average:**

It is a single numerical value which may be used to represent the whole data or distribution.

Method of finding arithmetic mean for ungrouped data.

$$\bar{x} = \frac{\sum x}{n} \quad (\text{Direct Method})$$

$$\bar{x} = A + \frac{\sum D}{n} \quad (\text{Short Formula})$$

where $D = x - A$

$$\bar{x} = A + \frac{\sum U}{n} \times h \quad (\text{Coding Method})$$

where $U = \frac{x - A}{h}$

Method of finding Arithmetic mean for grouped data.

$$\bar{x} = \frac{\sum fx}{\sum f} \quad (\text{Direct Method})$$

$$\bar{x} = A + \frac{\sum fd}{\sum f} \quad (\text{Short formula})$$

$$\bar{x} = A + \frac{\sum fu}{\sum f} \times h \quad (\text{Coding Method})$$

Q. 9 Write empirical relationship among mean median and mode.

Ans: Normally the relationship among above said three averages depend on the shape of given data or distribution.

i) For symmetrical data/distribution mean, median and mode coincide

$$\text{Mean} = \text{Median} = \text{Mode.}$$

ii) For moderately skewed data / distribution

$$\text{Mode} = 3 \text{ Median} - 2 \text{ Mean.}$$

Q.10 Write difference between primary and secondary data.

Ans: Primary data is raw data without any statistical treatment, while the secondary data is not raw-data and is available with statistical treatment.

Q.11 Classify the following variables as discrete or continuous.

Ans: i) Number of houses in a city
(Discrete)

ii) The yearly income of a shop keeper
(Continuous)

iii) The number of flower on a tree
(Discrete)

iv) The life time of electric bulbs
(Continuous)

v) The length of roll of cloth
(Continuous)

vi) Temperature of a patient
(Continuous)

Q.12 Briefly describe, why do we summarize the data?

Ans: It is important to note that the collected data provide basis to draw conclusion and make decisions. Before a programme is prepared the relevant data are collected after the collection of data, the next step is to present the data in a way that provides a clear picture about data thus obtained.

EXERCISE 4.1

Q.1 The following distribution represents the scores achieved by a group of science students in the chemistry laboratory. Answer the following questions.

SCORES	NO. OF STUDENTS
24-28	3
29-33	6
34-38	12
39-43	23
44-48	15
49-53	6
Total	65

- Q.** What is the upper limit of the last class?
 ii. What is the lower limit of the interval (39-43)?
 iii. What is the mid value in the interval (34-38)?

- iv. What are the class frequencies of the classes (29-33) and (44-48)?
 v. What is the size of the class interval in above frequency distribution?
 vi. In which class or group minimum number of students falls?
 vii. What is the lower limit of the class having 15 as its class frequency?
 viii. What is the number of students having scores between 24 and 43?

Sol: i. 53
 ii. 39
 iii. 36
 iv. 6 and 15
 v. 5
 vi. 24-48
 vii. 44
 viii. $3 + 6 + 12 + 23 = 44$

Q.2 Following are the mistakes made by a group of students of class 10th in a test of essay writing. Using an appropriate size of class interval, make frequency distribution and also indicate the number of class intervals.

4, 7, 12, 9, 21, 16, 3, 19, 17, 24, 14, 15, 8, 13, 11, 16, 15, 6, 5, 8, 11, 20, 18, 22, 6.

Sol: 3, 4, 5, 6, 6, 7, 8, 8, 9, 11, 11, 12, 13, 14, 15, 15, 16, 16, 17, 18, 19, 20, 21, 22, 24.

Largest number = 24

Smallest number = 3

Size of class interval = 3

Hence req. frequency distribution is

C. I	Tally marks	Frequency
3-5	III	3
6-8	THH	5
9-11	III	3
12-14	III	3
15-17	THH	5
18-20	III	3
21-23	II	2
24-26	I	1
Total		25

Q.3 For a school staff, the following expenditure (in rupees) is required for

Q.4 Make a frequency distribution using 4 as the size of class interval for the data given in question - 3 by tally method. Also state the class boundaries.

Sol: Largest number = 165

Here smallest number = 144

Size of class interval = 4

Class Interval	Tally Methods	Frequency	Class boundaries	Mid pt.
144-147	THH I	6	143.5-147.5	145.5
148-151	THH	5	147.5-151.5	149.5
152-155	II	9	151.5-155.5	153.5
156-159	THH	5	155.5-159.5	157.5
160-163	III	4	159.5-163.5	161.5
164-167	I	1	163.5-167.5	165.5
Total		30		

the repair of chairs 145, 152, 153, 156, 158, 160, 146, 152, 155, 159, 161, 163, 164, 165, 147, 148, 151, 154, 156, 158, 160, 144, 147, 151, 150, 152, 149, 145, 153, 152, 155.

Prepare a frequency distribution by (tally bar method) using 3 as the size of class interval and also. Write down what are the frequencies of the last three classes?

Sol: Largest numbers = 165

Smallest number = 144

Size of class interval = 3

Required frequency distribution is

C. I	Tally marks	Frequency
144-146	III	4
147-149	III	4
150-152	THH II	7
153-155	THH	5
156-158	III	4
159-161	III	4
162-164	I	1
165-167	I	1
Total		30

Now

Sum of frequencies of Last three

classes = 1 + 1 + 4

= 6

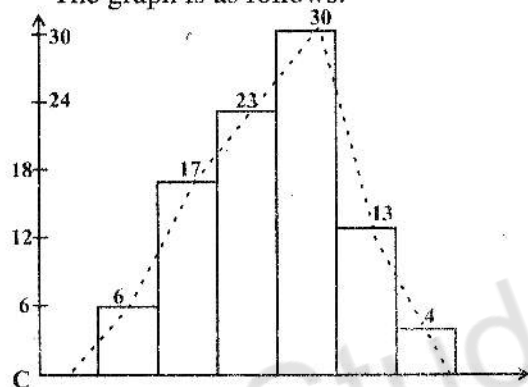
Q.5 Make a histogram of the frequency distribution obtained in question – 4.

Sol: The graph depends on the following step.

Draw two lines, horizontally and vertically on a graph paper.

ii. Select a suitable scale.

The graph is as follows:



Q.6 Given below are the weights in (lbs) of 30 students of a high school 130, 133, 124, 121, 115, 139, 137, 144, 142, 133, 133, 128, 129, 132, 131, 128, 126, 132, 134, 135, 138, 136, 141, 123, 126, 118, 134.

Take 5 as size of class interval, prepare a frequency table.

Sol: Highest Weight = 144 lbs

Lowest Weight = 115 lbs

Size of class interval = 5

The required frequency table is

Class Interval	Tally	Frequency
115-119	II	2
120-124	III	3
125-129	III	5
130-134	IIII IIII	10
135-139	IIII I	6
140-144	IIII	4
Total		30

The frequency is 2, 3, 4, 5, 10, 6, 4

Q.7 A group of 10th class students obtained the following marks (out of 100 marks) in English test.

58, 59, 58, 33, 40, 58, 45, 46, 43, 45, 45, 50, 52, 49, 50, 57, 63, 55, 49, 50, 65, 49, 48, 44, 42, 46, 42, 46, 53, 40, 44

Classify the data into a frequency distribution by (direct method) taking 6 as the size of class interval. Also find the class interval with least class frequency.

Sol: Highest Marks = 65

Lowest Marks = 33

Difference = 65 - 33 = 32

Size of class interval = 6

The frequency distribution table (by direct method) is as follows:

CLASS INTERVAL	MARKS	FREQUENCY
33-38	33	1
39-44	40, 40, 42, 43, 44, 44	6
45-50	45, 46, 45, 45, 50, 49, 50, 49, 50, 49, 50, 49, 48, 46, 46	15
51-56	52, 55, 53	3
57-62	58, 59, 58, 58, 57	5
63-69	63, 65	2
Total		32

The frequencies are 1, 6, 15, 3, 5, 2

The least class interval with least class frequency is (33-38)

Q.8 Following are the marks (out of 500) obtained by 40 students in a certain examination

310,350,370,320,380,390,400,398,399,315,326,337,348,368,359,361,372,382,389,309,340,335,301,302,317,345,350,335,354,340,400,365,335,375,342,332,325,376,374,338

Mark a frequency distribution taking 15 as size of the class interval and find the number of students having marks between 334 and 365 inclusive

Sol: Highest Marks = 400
Lowest Marks = 301

Size of Class interval = 15

Class Interval	Tally	No. of Students
301-315		5
316-330		4
331-345		10
346-360		5
361-375		7
376-390		5
391-405		4
Total		40

The frequencies are 5, 4, 10, 5, 7, 5, 4

The number of students having marks between 334 and 365 = 15

Q.9 From the Table given below, indicate which of the following statements are true or false?

Weights (Lbs.)	No. of Students	Mid Points
110-114	06	112
115-119	17	117
120-124	23	122
125-129	30	127
130-134	22	132
135-139	13	137
140-144	04	142

The lower limit in the class (115-119) is 120.

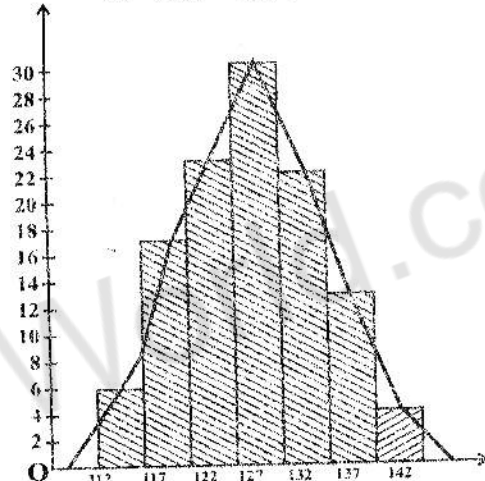
ii. The upper limit 129 is included in (125-129).

iii. The lower limit 135 is not included in (135-139).

iv. The no. of students having weights between (110-144) Lbs. is 1050.

Draw a frequency polygon from the above frequency distribution.

Sol: i. False, lower limit is 115
ii. True iii. False iv. False



Q.10 Following data shows the number of heads in an experiment of 50 sets of tossing a coin 5 times. Make a discrete frequency distribution from the information

3, 3, 4, 0, 5, 4, 3, 3, 1, 2, 4, 5, 0, 3, 2, 4, 4, 4, 0, 0, 0, 5, 5, 3, 2, 4, 3, 2, 5, 3, 2, 1, 2, 2, 1, 3, 5, 4, 3, 5, 4, 3, 2, 1, 3, 2, 1, 3, 2, 1, 3, 1, 3, 1, 4, 3, 2.

Sol: As the no. of heads in the experiment for tossing a coin 5 times is discrete, therefore we make a discrete frequency distribution

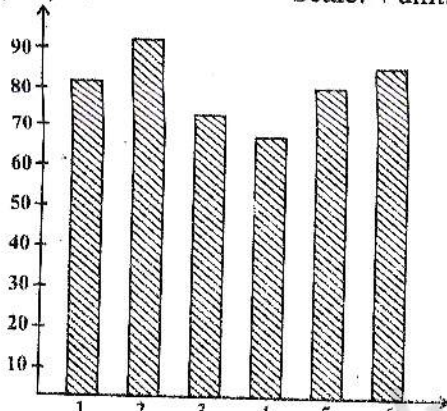
Heads	Numbers	Frequency
0	0,0,0,0,0	5
1	1,1,1,1,1,1,1	7
2	2,2,2,2,2,2,2,2,2,2	10
3	3,3,3,3,3,3,3,3,3,3,3,3,3,3	14
4	4,4,4,4,4,4,4,4,4	8
5	5,5,5,5,5,5	6
Total		50

Exercise 4.2

Q.1 The grades of a student "Areeba" in six examinations were 82, 91, 74, 68, 80 and 86. Show the above figure by making a simple bar diagram.

Sol: The marks obtained are 82, 91, 74, 68, 80, 86

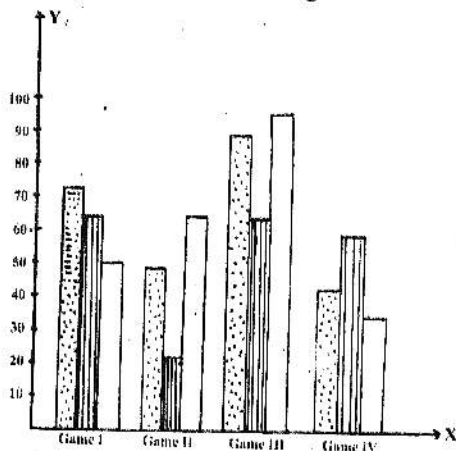
Scale: 4 units



Q.2 Prepare a multiple bar diagram to show the paints out of (100) made by four students in three games.

Students	Paints in Game I	Paints in Game II	Paints in Game III
Tahir	72	64	50
Adeel	48	22	64
Saad	88	62	94
Ashan	42	57	33

Sol: Req. multiple bar diagram is



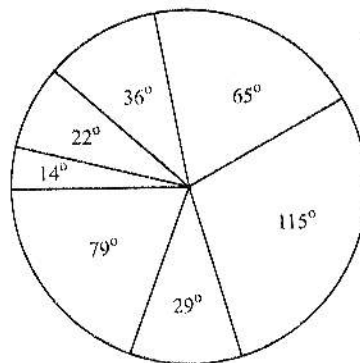
Q.3 The data given below shows the number of teachers according to their academic qualifications in an English medium school. Construct a pie-chart

Academic Qualifications	No. of Teachers
F.A.C.T	3
F.Sc.	5
B.A	9
B.A. B.Ed.	16
B.Sc.	4
B.Sc.B.Ed.	11
M.A / M.Sc.	2

Sol:

Academic Qualifications	No. of Teachers	Angle of Sector
F.A.C.T	3	$\frac{3}{50} \times 360 = 22^\circ$
F.Sc.	5	$\frac{5}{50} \times 360 = 36^\circ$
B.A	9	$\frac{9}{50} \times 360 = 65^\circ$
B.A B.Ed.	16	$\frac{16}{50} \times 360 = 115^\circ$
B.Sc	4	$\frac{4}{50} \times 360 = 29^\circ$
B.S.c. B.Ed.	11	$\frac{11}{50} \times 360 = 79^\circ$
M.A / M.Sc.	2	$\frac{2}{50} \times 360 = 14^\circ$
Total	50	360°

Pie-diagram is as follows:

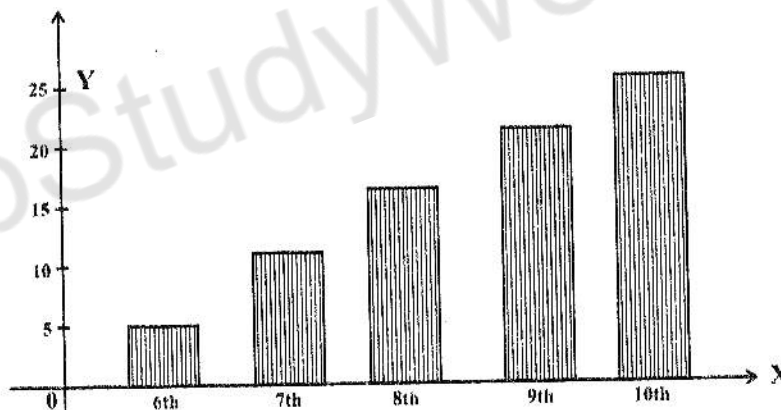


Q.4 A student sharique received the following number of prizes in his five classes.

Represent the above information by a simple bar diagram

SOL: Req. simple bar diagram is

Class	6 th	7 th	8 th	9 th	10 th
Prizes	5	12	17	21	25

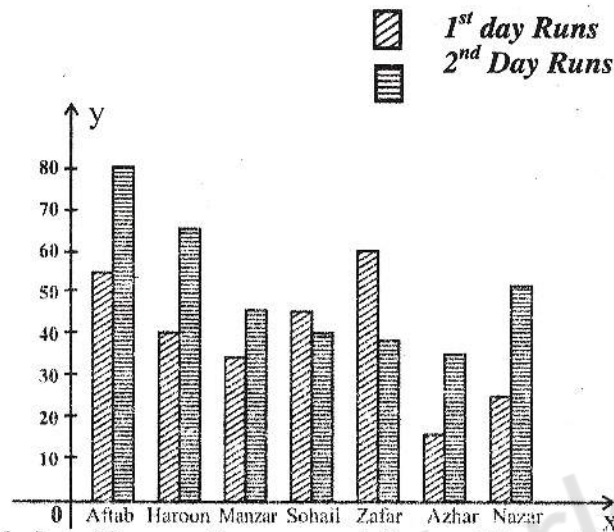


Q.5 The following are the runs made in a two days cricket match by the students in a high school.

Students	Runs Day 1	Runs Day II
Aftab	55	80
Haroon	40	66
Manzar	35	45
Sohail	45	40
Zafar	60	38
Azhar	16	35
Nazar	25	52

Draw a suitable diagram for the data given above

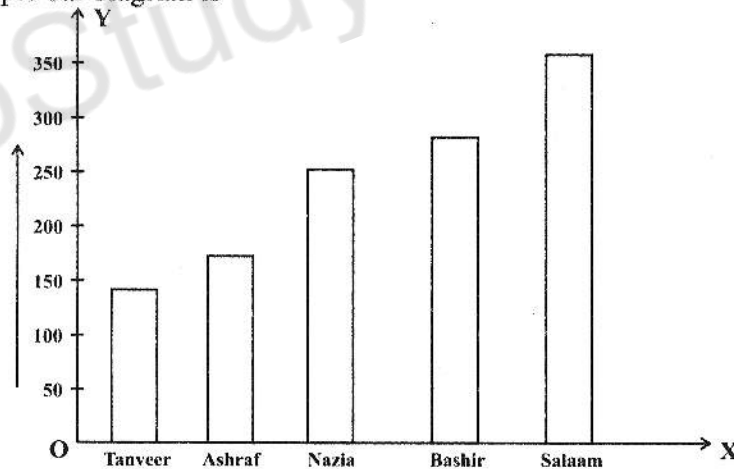
Sol:



Q.6 Draw a simple bar diagram to represent the weekly pocket money of s students of an English medium school

Students	Tanveer	Ashraf	Nazir	Bashir	Saleem
Pocket money	140	175	250	275	350

Sol: Req. simple bar diagram is



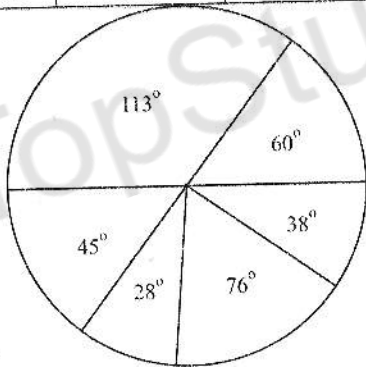
Q.7 A female house hold Mrs. Nighat brought the following goods and services from a general store for her kitchen in he last month.

Goods and Services	Experience (Rs.)
Flour	240
Rice	450
Dal	180
Sugar	110
Oil	300
Miscellaneous	150

Graph the data by a pie-diagram

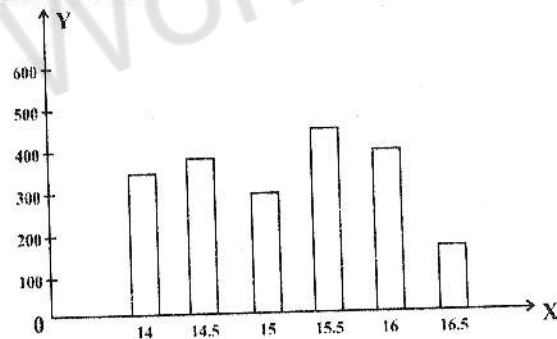
Sol:

Goods and Services	Expenditure (Rs.)	Angle of Sector
Flour	240	$\frac{240}{1430} \times 360 = 60^\circ$
Rice	450	$\frac{450}{1430} \times 360 = 113^\circ$
Dal	180	$\frac{180}{1430} \times 360 = 45^\circ$
Sugar	110	$\frac{110}{1430} \times 360 = 28^\circ$
Oil	300	$\frac{300}{1430} \times 360 = 76^\circ$
Miscellaneous	150	$\frac{150}{1430} \times 360 = 38^\circ$
Total	1430	360°



Q.8 A shirt making company makes the following number of shirts against their collar sizes in (inches) as given below. Show the data by a bar diagram.

Collar Sizes in (Inches)	No. of Shirts
14	350
$14\frac{1}{2}$	375
15	300
$15\frac{1}{2}$	450
16	400
$16\frac{1}{2}$	175



Exercise 4.3

Q.1 What is an average? Explain the method of calculation of arithmetic mean for ungrouped data as well as for grouped data.

Sol: AVERAGE OR ARITHMETIC MEAN

It is defined as a value which is obtained by dividing the sum of all the values by their number of observations thus A.M. of values

$$x_1, x_2, \dots, x_n \text{ is } \bar{x}$$

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$= \frac{\sum x_i}{n}$$

Where \sum stands for the sum and n is number of observations.

CALCULATION OF A.M. BY DIRECT METHOD

$$\bar{x} = \frac{\sum x}{n} \quad (\text{Ungrouped data})$$

$$\bar{x} = \frac{\sum fx}{\sum f} \quad (\text{Grouped data})$$

METHOD OF FINDING A.M. BY SHORT FORMULA METHOD:

$$\bar{x} = A + \frac{\sum D}{n}$$

(Ungrouped data)

Where $D = X - A$

$$\bar{x} = A + \frac{\sum fD}{\sum f}$$

(Grouped data)

CALCULATION OF A.M. BY CODLING METHOD

$$\bar{x} = A + \frac{\sum U}{n} \times h$$

(Ungrouped data)

$$\bar{x} = A + \frac{\sum fU}{\sum f} \times h$$

(Grouped data)

Q.2 Find the A.M. in each of the following

i. 4, 6, 10, 12, 15, 20, 25, 28, 30

SOL: $\bar{x} = \frac{\sum x}{n}$

$$\bar{x} = \frac{4+6+10+12+15+20+25+28+30}{9}$$

$$= \frac{150}{9}$$

$$= 16.6$$

ii. 12, 18, 19, 0, -19, -18, -12

SOL: $\bar{x} = \frac{\sum x}{n}$

$$\bar{x} = \frac{12+18+19+0-19-18-12}{7}$$

$$= \frac{0}{7}$$

$$= 0$$

iii. 6.5, 11, 12.3, 9, 8.1, 16, 18, 20.5, 25

Sol: $\bar{x} = \frac{\sum x}{n}$

$$\bar{x} = \frac{6.5+11+12.3+9+8.1+16+18+20.5+25}{9}$$

$$= \frac{126.4}{9}$$

$$= 14.04$$

iv. 8, 10, 12, 14, 16, 20, 22

Sol: $\bar{x} = \frac{\sum x}{n}$

$$\bar{x} = \frac{8+10+12+14+16+20+22}{7}$$

$$= \frac{102}{7}$$

$$= 14.57$$

Q.3 Define median and give its merits and demerits. How do we calculate median for grouped data?

SOL:

Median: When the data are arranged in some order. Median is the value which divides the data into two equal parts i.e. 50% of the data are before the median and 50% data after it.

Median of n observations x_1, x_2, \dots, x_n can be obtained as:

When n is odd number

$$\text{Median} = \left(\frac{n+1}{2} \right) \text{th item}$$

When n is even number

$$\text{Median} = \frac{1}{2} \left[\left(\frac{n}{2} \right) \text{th item} + \left(\frac{n+2}{2} \right) \text{th item} \right]$$

MEDIAN FOR GROUPED DATA

It is obtained by the following formula

$$\text{Median} = l + \frac{n}{f} \left(\frac{n}{2} - C \right)$$

Where

l = Lower class boundary of median class

h = Length of class interval of median class

f = Frequency of median class
 n = Total frequency i.e $\sum f$
 c = Cumulative frequency preceding the median class

Q.4 Following are the heights in (inches) of 12 students. Find the median height.

55, 53, 54, 58, 60, 61, 62, 56, 52, 51, 63, 57

Sol:

Arranging the heights of the students in ascending order

$x = 51, 52, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63$

Since n is even

$$\begin{aligned}\text{Median} &= \frac{1}{2} \left[\frac{n}{2} \text{th item} + \left(\frac{n+2}{2} \right) \text{th item} \right] \\ &= \frac{1}{2} \left[\frac{12}{2} \text{th item} + \left(\frac{12+2}{2} \right) \text{th item} \right] \\ &= \frac{1}{2} [6\text{th item} + 7\text{th item}] \\ &= \frac{1}{2} [56 + 57] \\ &= \frac{1}{2} [113] \\ &= 56.5\end{aligned}$$

Q.5 Define mode and explain the method of its calculation.

Sol:

Mode: in a data the value which occurs most often is called mode. It is the most common value

For grouped data, mode can be calculated by the following formula

$$\text{Mode} = l + \frac{(fm - f_1)}{(fm - f_1) + (fm - f_2)} \times h$$

Where

l = Lower class boundary of model class

fm = Frequency of model class

f_1 = Frequency preceding model class

f_2 = Frequency following model class
 h = Class interval of model class

Mode for ungrouped data

75, 76, 80, 82, 82, 83, 85

As 82 repeated more than any other numbers, early mode is 82.

Q.6 Following are the earnings (in Rs.) of ten workers

88, 70, 72, 125, 115, 95, 81, 90, 95, 90

Calculate

- Arithmetic mean
- Median
- Mode

Sol: $\bar{x} = \frac{\sum x}{n}$

A.M

$$\begin{aligned}&= \frac{88 + 70 + 72 + 125 + 115 + 95 + 81 + 90 + 95 + 90}{10} \\ &= \frac{921}{10} \\ &= 92.1\end{aligned}$$

ii. Median

Arrange the data as

70, 72, 81, 88, 90, 90, 95, 95, 115, 125

= 10

Since n is even, so

$$\begin{aligned}\text{Median} &= \frac{1}{2} \left[\frac{n}{2} \text{th item} + \left(\frac{n+2}{2} \right) \text{th item} \right] \\ &= \frac{1}{2} \left[\frac{10}{2} \text{th item} + \left(\frac{10+2}{2} \right) \text{th item} \right] \\ &= \frac{1}{2} [5\text{th item} + 6\text{th item}] \\ &= \frac{1}{2} [90 + 90] \\ &= \frac{1}{2} [180] \\ &= 90\end{aligned}$$

iii. Mode:

90, 95

Q.7 Marks obtained by the students in the subject of English are given below. Find

(i) A.M. of their marks

ii. Median of their marks

MARKS OBTAINED	STUDENTS
15-19	9
20-24	18
25-29	35
30-34	17
35-39	5

i. A.M

SOL:

Marks Obtained	Mid Pt. (x)	No. of students	fx
15-19	$\frac{15+19}{2} = 17$	9	$9 \times 17 = 153$
20-24	$\frac{20+24}{2} = 22$	18	$18 \times 22 = 396$
25-29	$\frac{25+29}{2} = 27$	35	$35 \times 27 = 945$
30-34	$\frac{30+34}{2} = 32$	17	$17 \times 32 = 544$
35-39	$\frac{35+39}{2} = 37$	5	$5 \times 37 = 185$
Total	-----	$\sum f = 84$	$\sum fx = 2223$

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{2223}{84} = 26.46$$

ii) Median

Marks Obtained	No. of students (f)	Class Boundaries (C.B)	Class frequency (C.F.)
15-19	9	14.5-19.5	9
20-24	18	19.5-24.5	$18+9=27 \rightarrow C$
25-29	35	24.5, 29.5	$35+27=62 \rightarrow \frac{n}{2}$
30-34	17	29.5-34.5	$17+62=79$
35-39	5	34.5-39.5	$5+79=84$
Total	$\sum f = 84$	-----	-----

$$\frac{n}{2} \text{th value} = \frac{84}{2} = 42$$

$$\text{Median} = l + \frac{h}{f} \left(\frac{n}{2} - C \right)$$

$$= 24.5 + \frac{5}{35} (42-27)$$

$$= 24.5 + \frac{1}{7} (42-27)$$

$$= 24.5 + \frac{1}{7} (15)$$

$$= 24.5 + 2.14$$

$$\text{Median} = 26.64$$

Q.8 Given below is a frequency distribution. Find mode of the frequency distribution.

Class Interval	(5-9)	(10-14)	(15-19)	(20-24)	(25-29)
Frequency	1	8	18	11	2

Class Interval	(f)	C.b
5 - 9	1	4.5-9.5
10 - 14	8 $\rightarrow f_1$	9.5-14.5
15 - 19	18 $\rightarrow f_m$	14.5-19.5
20 - 24	11 $\rightarrow f_2$	19.5-24.5
25 - 29	2	24.5-29.5

SOL: Mode:

$$\begin{aligned} \text{Model} &= l + \frac{(f_m - f_1)}{(f_m - f_1)(f_m - f_2)} \times h \\ &= 14.5 + \frac{(18-8) \times 4}{(18-8) + (18-11)} \\ &= 14.5 + \frac{10 \times 4}{10+7} \\ &= 14.5 + \frac{40}{17} \\ &= 17.44 \end{aligned}$$

Q.9 Ten boys work on a petrol pump station. They get weekly wages as under. Wages in (Rs.) 425, 435, 440, 425, 435, 441, 450, 430, 450, 439

Find arithmetic mean by short formula, Median and mode of their wages

SOL:

i. Arithmetic mean

$$\begin{aligned} \bar{x} &= \frac{425 + 435 + 440 + 425 + 435 + 441 + 450 + 430 + 450 + 439}{10} \\ &= \frac{4370}{10} \\ \bar{x} &= 437 \end{aligned}$$

ii. Median

425, 425, 430, 435, 435, 439, 440, 441, 450, 450

Since n is even, so

$$\begin{aligned} &= \frac{1}{2} \left[\frac{n}{2} \text{th item} + \left(\frac{n+2}{2} \right) \text{th item} \right] \\ &= \frac{1}{2} [5\text{th item} + 6\text{th item}] \\ &= \frac{1}{2} [435 + 439] \\ &= \frac{1}{2} [874] \\ &= \frac{874}{2} \\ &= 437 \end{aligned}$$

iii. Mode

Mode = 425, 435, 450

Q.10 The arithmetic mean of 45 numbers is 80. Find the sum

SOL: Arithmetic mean $\bar{x} = 80$

$$n = 45$$

$$\sum x = ?$$

$$\text{As } \bar{x} = \frac{\sum x}{n}$$

$$\begin{aligned} \sum x &= n \bar{x} \\ &= 45 \times 80 \\ &= 3600 \end{aligned}$$

Q.11 Five numbers are 1, 4, 0, 7, 9 find its mean, median and mode

SOL:

$$\text{i. A.M} = \bar{x} = \frac{1+4+0+7+9}{5}$$

$$= \frac{21}{5}$$

$$= 4.2$$

ii. Median

$$x = 1, 4, 0, 7, 9$$

Writing in ascending order

$$x = 0, 1, 4, 7, 9$$

Since n is odd, so

$$\text{Median} = \left(\frac{n+1}{2}\right)\text{th term}$$

$$= \frac{5+1}{2}\text{th term}$$

$$= 3^{\text{th}} \text{ term}$$

$$\text{Median} = 4$$

iii. Mode

Since no term appears more than once so it has no mode

Q.12 A set of data contains the values as 148, 145, 160, 157, 156, 160. Show that $\text{mode} > \text{Median} > \text{Mean}$

SOL:

First we will find mean

$$\bar{x} (\text{Mean}) = \frac{148+145+160+157+160}{6}$$

$$= \frac{926}{6}$$

$$= 154.33$$

ii. Median

Writing the terms in ascending order

145, 148, 156, 157, 160, 160

Since n is even, so

$$\text{Median} = \frac{1}{2} \left[\left(\frac{n}{2}\right)\text{th item} + \left(\frac{n+2}{2}\right)\text{th item} \right]$$

$$= \frac{1}{2} \left[\left(\frac{6}{2}\right)\text{th item} + \left(\frac{6+2}{2}\right)\text{th item} \right]$$

$$= \frac{1}{2} [3\text{th item} + 4\text{th item}]$$

$$\text{Median} = \frac{1}{2} [3\text{th item} + 4\text{th item}]$$

$$= \frac{1}{2} [156 + 157]$$

$$= \frac{1}{2} [313]$$

$$= 156.5$$

iii. Mode

$$\bar{x} = 145, 148, 156, 157, 160, 160$$

Since in the data, 160 appears twice therefore mode is 160

From (i), (ii) and (iii)

$\text{Mode} > \text{Median} > \text{Mean}$

$$160 > 156.5 > 154.5$$

Q.13 The monthly attendance of 10 students for their lunch in the hostel is recorded as

21, 15, 16, 18, 14, 17, 15, 12, 13, 11

Find Median and Mode of attendance,

Also find mean if $D = x - 20$

*** SOL:**

Arrange the attendance in ascending order

11, 12, 13, 14, 15, 16, 17, 18, 21

Since $n = 10$ is even, so

$$= \frac{1}{2} \left[\left(\frac{n}{2}\right)\text{th item} + \left(\frac{n+2}{2}\right)\text{th item} \right]$$

$$= \frac{1}{2} \left[\left(\frac{10}{2}\right)\text{th item} + \left(\frac{10+2}{2}\right)\text{th item} \right]$$

$$\text{Median} = \frac{1}{2} [5^{\text{th}} \text{ item} + 6^{\text{th}} \text{ item}]$$

$$= \frac{1}{2} [15 + 15]$$

$$= \frac{1}{2} (30)$$

$$= 15$$

$\text{Mode} = 11, 12, 13, 14, 15, 15, 16, 17, 18, 21$

Since 15 appears 2 items in the data therefore mode is 15.

X	D = X - A
21	1
15	-5
16	-4
18	-2
14	-6
17	-3
15	-5
12	-8
13	-7
11	-9
$\sum D$	-48

$$\bar{x} = A + \frac{\sum D}{n}$$

$$\bar{x} = 20 + \left(-\frac{48}{10}\right)$$

$$\bar{x} = 20 - 4.8$$

$$\bar{x} = 15.2$$

Q.14 On a prize distribution day, 50 students brought pocket money as under:

RUPEES	NO. OF STUDENTS
5-10	12
10-15	9
15-20	18
20-25	7
25-30	4

- Find Median and Mode of above data
- Find arithmetic mean of data using coding method

SOL:

(i) Median = $l + \frac{h}{f} \left(\frac{n}{2} - c \right)$

Rupees	F	C.F
5-10	12	12
10-15	9	9+12=21 → C
ℓ 15-20	18 f	18+21=39 → $\frac{n}{2}$
20-25	7	7+39=46
25-30	4	4+46=50
Total	$\sum f = 50$	

$$\text{Median} = l + \frac{h}{f} \left(\frac{n}{2} - c \right)$$

$$= 15 + \frac{5}{18} (25-21)$$

$$\text{Median} = 15 + \frac{5}{18} (4)$$

$$= 15 + 1.1$$

$$= 16.11$$

(ii) Mode

Rupees	No. of students
5-10	12
10-15	9 → f_1
15-20	18 → f_m
20-25	7 → f_2
25-30	4

$$\text{Mode} = l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 15 + \frac{(18-9) \times 5}{(18-9) + (18-7)}$$

$$= 15 + \frac{9 \times 5}{9+11}$$

$$= 15 + \frac{45}{20}$$

$$= 15 + 2.25$$

$$= 17.25$$

$$\text{Mode} = 17.25$$

ii. Now we will find mean of the distribution

Rupees	Mid Pt. (X)	No. of students	$U = \frac{X - A}{h}$	fU
5-10	$\frac{5+10}{2} = 7.5$	12	-2.5	-30
10-15	$\frac{10+15}{2} = 12.5$	9	-1.5	-13.5
15-20	$\frac{15+20}{2} = 17.5$	18	-0.5	-9
20-25	$\frac{20+25}{2} = 22.5$	7	0.5	3.5
25-30	$\frac{25+30}{2} = 27.5$	4	1.5	6
Total		$\sum f = 50$		$\sum fU = -43$

$$\begin{aligned} \text{Arithmetic mean} &= A + \frac{\sum fU}{\sum f} \times h \\ &= 20 + \frac{-43}{50} \times 5 \\ &= 20 - 4.3 \\ &= 15.7 \end{aligned}$$

Q.15 The arithmetic mean of the ages of 20 boys is 13 years, 4 months and 5 days. Find the sum of their ages if one of the boy is of age exactly 15 years. What is the average age of the remaining boys?

SOL:

(i) Given $n = 20$

$\bar{x} = 13$ years, 4 months and 5 days

$$\begin{aligned} &= 13 + \frac{4}{12} + \frac{5}{12 \times 30} \\ &= 13 + \frac{1}{3} + \frac{1}{72} \\ &= \frac{936 + 24 + 1}{72} \\ &= \frac{961}{72} \text{ years} \end{aligned}$$

$$\bar{x} = \frac{\sum x}{n}$$

$$961 = \frac{\sum x}{20}$$

$$\sum x = 20 \times \frac{961}{72}$$

$$= 266.944 \text{ year}$$

$$= 266 \text{ years, 11 months and 10 days.}$$

ii- Sum of ages of 20 boys = 266.944 years

age of 1 boy = 15 years

Sum of ages of 19 boys = 266.944 - 15

$$\text{Average age} = \frac{251.944}{19}$$

$$= 13.26 \text{ years}$$

$$= 13 \text{ years, 3 months and 4 days.}$$

So req. average age is 13 years, 3 months, 4 days approx.

Q.16 Calculate arithmetic mean from the following information

i. If $D = X - 140$, $\sum D = 500$,

$n = 10$

SOL:

Since $D = X - A$

So $A = 140$

Hence

$$\bar{x} = A + \frac{\sum D}{n}$$

$$= 140 + \frac{500}{10}$$

$$= 140 + 50$$

$$= 190$$

ii. If $u = \frac{x-130}{6}$, $\sum u = -150$, $n = 15$

SOL: Hence $A = 130$

and $h = 6$

$$\bar{x} = A + \frac{\sum u}{n} \times h$$

$$= 130 - \frac{150}{15} \times 6$$

$$= 130 - 10 \times 6$$

$$= 130 - 60$$

$$= 70$$

iii. $D = X - 25$, $\sum fD = 300$ and $\sum f = 20$

SOL:

Here $A = 25$

$$\bar{x} = A + \frac{\sum fD}{\sum f}$$

$$\bar{x} = 25 + \frac{300}{20}$$

$$= 25 + 15$$

$$\bar{x} = 40$$

iv.

If $u = \frac{x-120}{5}$, $\sum fu = 60$ and $\sum f = 100$

SOL:

Here $A = 120$

$h = 5$

So

$$\bar{x} = A + \frac{\sum fu}{\sum f} \times h$$

$$= 120 + \frac{60}{100} \times 5$$

$$= 120 + 3$$

$$\bar{x} = 123$$

Q.17 The three children Haris, Masab and Minal made the following scores in a game conducted by a group of teachers in the school.

HARIS SCORES	MASAB SCORES	MINAL SCORES
50	75	80
55	60	77
70	60	66
85	45	42
90	53	48

It is decided that candidate who gets highest average score will be awarded rupees 1000. Who will get the amount?

SOL:

Average scores of Haris =

$$\frac{50+55+70+85+90}{5}$$

$$= \frac{350}{5}$$

$$= 70$$

Average scores of Masab

$$= \frac{75+60+60+45+53}{5}$$

$$= \frac{293}{5}$$

$$= 58.6$$

Average scores of Minal =

$$\frac{80+77+66+42+48}{5}$$

$$= \frac{313}{5}$$

As average scores of Haris is highest so Haris will win the prize

Q.18 Given below is a frequency distribution derived by making a substitution as $D = X - 20$ Calculate the arithmetic mean

D	-6	-4	-2	0	2	4	6
f	1	3	6	20	26	12	2

SOL:

As $D = X - 20$

So $A = 20$

D	f	fD
-6	1	-6
-4	3	-12
-2	6	-12
0	20	0
2	26	52
4	12	48
6	2	12
	$\sum f = 70$	$\sum fD = 82$

$$\begin{aligned}\bar{x} &= A + \frac{\sum fD}{\sum f} \\ &= 20 + \frac{82}{70} \\ &= 20 + 1.17 \\ \bar{x} &= 21.17\end{aligned}$$

Q.19 Being partners pinky and Sarah took part in a Quiz programme. They made the following number of points 45, 51, 58, 61, 74, 48, 46 and 50. Compute the average numbers of points using $D = X - 58$

SOL:

X	D = X - 58
45	-13
51	-7
58	0
61	3
74	16
48	-10
46	-12
50	-8
\sum	$\sum D = -31$

Here $A = 58$

$n = 8, \sum D = -31$

$$\begin{aligned}\bar{x} &= A + \frac{\sum D}{n} \\ &= 58 + \frac{-31}{8} \\ &= 58 - 3.87 \\ \bar{x} &= 54.13\end{aligned}$$

Exercise 4.4

Q.1 What is range? How is it calculated?

SOL: Range: It is obtained by taking the difference between the largest and the smallest value in a given data it is calculated as:

Range = $x_m - x_l$

Where x_m = largest value

x_l = smallest value

Q.2 Find the range of the following set of measurements.

1.25, 3.12, 15, 14.28, 9.1, 16.24, 4.45, 1.13, 9, 16.28

SOL:

Here $R = x_m - x_l$

$x_m = 16.28$

$x_l = 1.13$

$R = x_m - x_l$

$= 16.28 - 1.13$

$R = 15.15$

Q.3 Define variance. Give a method for its calculation

SOL:

Variance

The variance is obtained by dividing the sum of the squares of the deviations taken

from arithmetic mean by the number of observations in the data set. It is denoted by S^2 and is defined for ungrouped data as:

$$i. S^2 = \frac{\sum (x - \bar{x})^2}{n}$$

(Proper mean Formula)

$$\text{or } S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2 \quad (\text{Direct formula})$$

Q.4 Find the variance of the following set of observations

i. $x=5, 13, 15, 25, 12, 18, 17, 19, 20, 16, 3$

SOL:

Here $x = 5, 13, 15, 25, 12, 18, 17, 19, 20, 16, 3$

X	x^2
5	25
13	169
15	225
25	625
12	144
18	324
17	289
19	361
20	400
16	256
3	9
$\sum x = 163$	$\sum x^2 = 2827$

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

$$= \frac{2827}{11} - \left(\frac{163}{11} \right)^2$$

$$= 257 - 219.57$$

$$= 37.42$$

$$S^2 = 37.42$$

ii. 110, 115, 108, 95, 118, 130, 112, 116, 120, 190

SOL:

X	X^2
110	12100
115	13225
108	11664
95	9025
118	13924
130	16900
112	12544
116	13456
120	14400
190	36100
1214	153338

Now

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

$$= \frac{153338}{10} - \left(\frac{1214}{10} \right)^2$$

$$= 15333.8 - 14737.96$$

$$= 595.84$$

$$S^2 = 595.84$$

Q.5 What is standard deviation? Describe the method of its calculation.

SOL: STANDARD DEVIATION

The standard deviation is the positive square root of variance. It is abbreviated as S.D. and is denoted by S. It is calculated by the following formula for ungrouped data.

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n}} \quad (\text{Proper mean formula})$$

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2} \quad (\text{Direct formula})$$

Q.6 Elaborate the dispersion in each of the following grades

i. $x = 1, 1, 0, 1, 1$

SOL:

i. $x = 1, 1, 0, 1, 1$

$$\text{Range} = x_m - x_l$$

$$= 1 - 0$$

$$= 1$$

STANDARD DEVIATION

x	x ²
1	1
1	1
0	0
1	1
1	1
$\sum x = 4$	$\sum x^2 = 4$

$$S.D = S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{4}{5} - \left(\frac{4}{5}\right)^2}$$

$$= \sqrt{\frac{4}{5} - \frac{16}{25}}$$

$$= \sqrt{\frac{20-16}{25}}$$

$$= \sqrt{\frac{4}{25}}$$

$$= \frac{2}{5}$$

$$S = 0.4$$

$$\begin{aligned} \text{(Variance)} S^2 &= (0.4)^2 \\ &= 0.16 \end{aligned}$$

ii. 4, 4, 4, 4, 4

SOL:

$$\text{Range} = x_m - x_l$$

$$= 4 - 4$$

$$= 0$$

Now we find S.D. and Variance

x	x ²
4	16
4	16
4	16
4	16
4	16
$\sum x = 20$	$\sum x^2 = 80$

$$S.D = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{80}{5} - \left(\frac{20}{5}\right)^2}$$

$$S.D = \sqrt{16 - (4)^2}$$

$$= \sqrt{16-16}$$

$$S = 0$$

Now

$$\text{Variance} = S^2 = (0)^2 = 0$$

Q.7 The following are the marks obtained out of 100 by eight students in two subjects

Students	1	2	3	4	5	6	7	8
Marks (Urdu)	45	42	33	38	47	51	35	60
Marks (Civics)	48	50	41	35	50	61	42	62

Compare the standard deviations of marks in the two subjects.

In which subject the students are more consistent?

SOL:

Marks in Urdu as:

45, 42, 33, 38, 47, 51, 35, 60

Students	x	x ²
1	45	2025
2	42	1764
3	33	1089
4	38	1444
5	47	2209
6	51	2601
7	35	1225
8	60	3600
	$\sum x = 351$	$\sum x^2 = 15957$

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{1994.625 - 1925.015}$$

$$= \sqrt{69.61}$$

$$S = 8.34$$

Now

$$\bar{x} = \frac{45+42+33+38+47+51+35+60}{8}$$

$$\bar{x} = \frac{351}{8}$$

$$\bar{x} = 43.88$$

$$\text{C.V for Urdu} = \frac{S}{\bar{x}} \times 100$$

$$= \frac{8.34}{43.88} \times 100$$

$$= 19 \%$$

Now Marks in civics are

48, 50, 41, 35, 50, 61, 42, 62

Students	y	y ²
1	48	2304
2	50	2500
3	41	1681
4	35	1225
5	50	2500
6	61	3721
7	42	1764
8	62	3844
Total	$\sum y = 389$	$\sum y^2 = 19539$

$$\sum y = 389$$

$$\sum y^2 = 19539$$

Now

$$S = \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2}$$

$$S = \sqrt{\frac{19539}{8} - \left(\frac{389}{8}\right)^2}$$

$$= \sqrt{2442.375 - 2364.39}$$

$$= \sqrt{77.9}$$

$$= 8.83$$

$$S = 8.83$$

Now

$$\bar{y} = \frac{48+50+41+35+50+61+42+62}{8}$$

$$= \frac{389}{8}$$

$$\bar{y} = 48.63$$

$$\text{C.V for civics} = \frac{S}{\bar{y}} \times 100$$

$$= \frac{8.83}{48.63} \times 100$$

$$= 18 \%$$

So in civics, the students are more consistent

Q.8 Calculate range and variance from the following marks (out of 100)

$\bar{x} = 46, 32, 18, 16, 14, 12, 60, 55, 52, 48, 40$

SOL:

$$\text{Range} = R = x_m - x_l$$

$$= 60 - 12$$

$$= 48$$

Number (x)	x ²
46	2116
32	1024
18	324
16	256
14	196
12	144
60	3600
55	3025
52	2704
48	2304
40	1600
$\sum x = 393$	$\sum x^2 = 17293$

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2$$

$$= \frac{17293}{11} - \left(\frac{293}{11}\right)^2$$

$$S^2 = 1572.09 - 1276.43$$

$$= 295.66$$

$$S^2 = 295.65$$

Q.9 The mean of numbers 4, 9, 3, 15, 19 is 10. Find the variance and standard deviation of the data.

SOL: Here $\bar{x} = 10$

x	x - \bar{x}	(x - \bar{x}) ²
4	-6	36
9	-1	1
3	-7	49
15	-5	25
19	-9	81
		192

$$\text{Now } S^2 = \frac{\sum (x - \bar{x})^2}{n}$$

$$= \frac{192}{5}$$

$$S^2 = 38.4$$

Now

$$S = \sqrt{38.4}$$

$$S = 6.20$$

Q.10 Find the variance and standard deviation from the following information.

$$\bar{x} = 19.5, \sum x^2 = 5555, \sum x = 195$$

SOL:

$$\text{Here } \bar{x} = 19.5$$

$$\sum x = 195$$

$$\sum x^2 = 5555$$

$$n = \frac{\sum x}{\bar{x}}$$

$$= \frac{195}{19.5}$$

$$n = 10$$

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

$$= \frac{5555}{10} - \left(\frac{195}{10} \right)^2$$

$$= 555.5 - 380.25$$

$$= 175.25$$

$$S^2 = 175.25$$

Now

$$S = \sqrt{175.25}$$

$$S = 13.238$$

Q.11 The height measurements (in inches) of 16 students for the selection of a game are given below:

64, 67, 62, 66, 63, 64, 63, 69, 63, 65, 67, 71, 65, 64, 72, 66

Calculate its mean and standard deviation. Find the percentage of observations lying within the limits

$$(\bar{x} \pm 1S), (\bar{x} \pm 2S), (\bar{x} \pm 3S)$$

SOL:

Here 64, 67, 62, 66, 63, 64, 63, 69, 63, 65, 67, 71, 65, 64, 72, 66

Now

x	x ²
64	4096
67	4489
62	3844
66	4336
63	3969
64	4096
63	3969
69	4761
63	3969
65	4225
67	4489
71	5041
65	4225
64	4096
72	5184
66	4356
1051	69165

$$\bar{x} = \frac{\sum x}{n} = \frac{1051}{16} = 65.69$$

Now

$$S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{69165}{16} - \left(\frac{1051}{16}\right)^2}$$

$$= \sqrt{4322.81 - 4314.84}$$

$$= \sqrt{7.97}$$

$$= 2.82$$

$$S = 2.82$$

Now we find limits ($\bar{x} \pm 1S$)

$$\text{Now } \bar{x} \pm 1S = (65.69 \pm 2.82)$$

$$= (62.87 \text{ and } 68.51)$$

There are 12 values which lying within the given limits

So %age of values lying within ($\bar{x} \pm 1S$)

$$= \frac{12}{16} \times 100$$

$$= 75\%$$

Now we will find limits ($\bar{x} \pm 2S$)

$$\bar{x} \pm 2S = (65.69 \pm 2(2.82))$$

$$= (65.69 \pm 5.64)$$

$$= (60.05 \text{ and } 71.33)$$

There are 15 values lying within given limits

So

%age of values lying within ($\bar{x} \pm 2S$)

$$= \frac{15}{16} \times 100$$

$$= 93.75\%$$

Now we find ($\bar{x} \pm 3S$)

$$\bar{x} \pm 3S = (65.69 \pm 3(2.82))$$

$$= (65.69 \pm 8.46)$$

$$= (57.23 \text{ and } 74.15)$$

All values be within the limits 57.23 and 71.15

% age of values lying within $\bar{X} \pm 3S$

$$= \frac{16}{16} \times 100$$

$$= 100\%$$

Q.1 Briefly describe, why do we summarize the data?

SOL: The collected data provide a basis to draw conclusions and make decisions. Before any programme is prepared. The relevant data are collected.

The Government makes budget for the coming year on the basis of the data of previous year by which we come to know that how much money the Government wants to spend on education, health, industry, transport etc. The teachers guess the ability of their students on the basis of their marks.

After collection of data, the next step is present the data in a way that provides a clear picture about the data thus obtained. The data in this way is easy to understand.

Q.2 What is a frequency distribution? Explain the steps involved in the construction of a frequency distribution.
SOL: Frequency Distribution:

A method of organizing and condensing the raw data is the conversion of raw data into a grouped data. The various items of data are classified into certain groups or classes and the number of items lying in each group or class is put against that group or class.

The data organized and summarized in this way is known as frequency distribution.

Q.3 Following are the weights of 40 students recorded to the nearest (lbs.)

138, 164, 150, 132, 144, 125, 149, 157,
146, 158, 140, 147, 136, 148, 152, 144,
168, 126, 138, 176, 163, 119, 154, 165,

146, 173, 142, 147, 135, 153, 140, 135,
161, 145, 135, 142, 150, 156, 145, 128

Make a frequency table taking size of class interval as 9.

SOL: The frequency table of class interval as 9 is

Weights (Lbs.)	No. of students	Class Boundaries	Mid pt. (x)
119-127	3	118.5-127.5	123
128-136	6	127.5-136.5	132
137-145	10	136.5-145.5	141
146-154	11	145.5-154.5	150
155-163	5	154.5-163.5	159
164-172	3	163.5-172.5	168
173-181	2	172.5-181.5	177
Total	40		

Q.4 Given below are the marks obtained by 45 students in Biology at a monthly test.

Marks	20-24	25-29	30-34	35-39	40-44	45-49
Students	5	8	12	15	3	2

With reference to the above table, find the following

- Upper class boundary of class 5th.
- Lower class boundary of last class
- Mid point of the class 2nd.
- Size of class interval
- Frequency of class 4th.
- The class interval with least frequency

SOL:

Marks	Class boundaries	f
20 – 24	19.5 – 24.5	5
25 – 29	24.5 – 29.5	8
30 – 34	29.5 – 34.5	12
35 – 39	34.5 – 39.5	15
40 – 44	39.5 – 44.5	3
45 – 49	44.5 – 49.5	2

- Upper class boundary of the class is 44.5
- Lower class boundary of last class is 44.5
- Mid pt. of 2nd class is 27
- Size of class interval is 5
- Frequency of fourth class is 15
- Class interval with least frequency is 45 – 49

Q.5 What is a measure of central tendency? briefly discuss its use. The grades of a student "Saad" in six tests were 84, 91, 72, 68, 87, 78.

Find arithmetic mean of his grades

SOL: MEASURES OF CENTRAL TENDENCY

The central value around which the data appeared to be crowded is known as a measure of central tendency. As it locates the centre of a distribution, so it is also known as a measure of location.

Following are the types of measure of central tendency

- Arithmetic mean
- Geometric mean
- Harmonic mean
- Median
- Mode

Now

SOL: A.M. by short method and coding method

Load	No. of roper (f)	Mid pt. (x)	D = x - 110	fD	U = $\frac{x-110}{5}$	fu
93-97	2	95	95-110 = -15	-30	-3	-6
98-102	5	100	-10	-50	-2	-10
103-107	8	105	-5	-40	-1	-8
108-112	12	110	0	0	0	0
113-117	6	115	5	30	1	6
118-122	2	120	10	20	2	4
	$\sum f = 35$	-----		-70		$\sum fu = -14$

$$\bar{x} = A + \frac{\sum fD}{\sum f} \text{ (Short cut-method)}$$

$$= 110 - \frac{70}{35}$$

$$= 110 - 2$$

$$\bar{x} = 108$$

Now we find \bar{x} by coding method

$$\bar{x} = A + \frac{\sum fu}{\sum f} \times h$$

$$\bar{x} = \frac{84+91+72+68+87+78}{6}$$

$$= \frac{480}{6}$$

$$= 80$$

$$\text{So } \bar{x} = 80$$

Q.6 Following distribution showed maximum load in (kg) supported by certain ropes. Find mean load using "short method" and coding method.

Max. load (kg)	No. of ropes
93-97	2
98-102	5
103-107	8
108-112	12
113-117	6
118-122	2

$$= 110 + \frac{-14}{35} \times 5$$

$$= 110 - 2$$

$$\bar{x} = 108$$

Q.7 Usman rolled two dice eight times. Each time their sum was recorded as 8, 5, 6, 6, 9, 4, 3, 11. Find the median and the mode of sum.

SOL: Writing the data in ascending order 3, 4, 5, 6, 6, 8, 9, 11

As n is even, so

Median

$$= \frac{1}{2} \left[\left(\frac{n}{2} \right) \text{th item} + \left(\frac{n+2}{2} \right) \text{th item} \right]$$

$$= \frac{1}{2} \left[\left(\frac{8}{2} \right) \text{th item} + \left(\frac{8+2}{2} \right) \text{th item} \right]$$

$$= \frac{1}{2} [4\text{th item} + 5\text{th item}]$$

$$= \frac{1}{2} [6 + 6]$$

Median = 6

Mode is 6

Q.8 Find the median and mode of the frequency distribution in Q-6. Also find the class mark i.e., the mid points of frequency distribution.

SOL:

Max. load kg	Frequency	C. B.	C.F
93-97	2	92.5-97.5	2
98-102	5	97.5-102.5	2 + 5 = 7
103-107	8	102.5-107.5	8 + 7 = 15 → C
108-112	12	107.5-112.5	12 + 15 = 27 → $\frac{n}{2}$
113-117	6	112.5-117.5	27 + 6 = 33
118-122	2	117.5-122.5	2 + 33 = 35
	$\sum f = 35$		

$$\text{Median} = l + \frac{h}{f} \left(\frac{n}{2} - C \right)$$

$$= 107.5 + \frac{5}{12} \left(\frac{35}{2} - 15 \right)$$

$$= 107.5 + \frac{5}{12} (2.5)$$

$$= 107.5 + 1.04$$

$$= 108.54$$

Now we find mode

Max. load (kg)	No. Of ropes
93-97	2
98-102	5
103-107	8 → f_1
108-112	12 → f_m
113-117	6 → f_2
118-122	2

Here $n = 5$

Mode

$$= l + \frac{(f_m - f_1)}{(f_m - f_1) + (f_m - f_2)} \times h$$

$$= 108 + \frac{(12 - 8) \times 5}{(12 - 8) + (12 - 6)}$$

$$= 108 + \frac{4 \times 5}{4 + 6}$$

$$= 108 + \frac{20}{10}$$

$$= 108 + 2$$

$$= 110$$

Q.9 For a set of 10 observations the mean is 25, later on checking, it was discovered that the observation 19 was incorrectly recorded where as the correct value was 23.

Calculate the correct mean from the information

SOL:

Mean of 10 observations = 25

Sum = 25 × 10 = 250

Correct Value = 23

Incorrect Value = 19

Difference = 23 - 19 = 4

Sum of 10 correct observations

$$= 250 + 4$$

$$= 254$$

So Correct mean value

$$= \frac{254}{10}$$

$$= 25.4$$

Q.10 Differentiate between discrete and continuous variable. Classify the following variables as discrete or continuous.

- The number of students in a class.
- Milk produced by a cow.
- The number of heads in tossing two coins
- The number of sixes in rolling 4 dice.
- The age of the employee of an office.
- The lifetime of T.V. tubes.
- The number of prize bond sold everyday.

Ans. Discrete Variable:

It can take only whole number e.g. if we toss coins, the number of heads is a discrete value as it can be one of the values 0, 1, 2, 3, 4, 5.

Continuous Variables:

It can take every possible value in a given interval (a,b). it may be a whole figure or a fraction. e.g. production of milk by a cow is a continuous variable.

Ans. (a) Discrete (b) continuous variable

(c) Discrete variable

(d) Discrete variable

(e) continuous variable

(f) continuous variable

(g) Discrete variable

Q.11 Explain the meaning of the term "Dispersion". What are the types of dispersion give one method of measuring dispersion.

SOL: DISPERSION

The measure in which the various values are spread out is known as measure of dispersion. It is possible that different

sets of data may have the same average but their dispersion around the central value may be different.

In fact, measures of variability provide an estimate about the dispersion. This measure indicates how the values tend to be scattered around the central value. The measure of dispersion play an important role in drawing conclusion and make decision in certain situation. To compare the variation of two or more series measured in defined units a measure of relative given by

$$C.V = \frac{S}{x} \times 100$$

Smaller coefficient of variation is more consistent in performance

Q.12 The heights of 11 player of a football team are 4'-9", 5'-1", 5'-0", 5'-4", 4'-11", 4'-7", 4'-10", 5'-3", 5'-5", 5'-1", 4'-8". Find the variance and S.D. of the heights of these players.

SOL: Note: 4' - 9" = 4 × 12 + 9 = 57

X	X ²
57	3249
61	3721
60	3600
64	4096
59	3481
55	3025
58	3364
63	3969
65	4225
61	3721
56	3136
$\sum x = 659$	$\sum x^2 = 39587$

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

$$= \frac{39587}{11} - \left(\frac{659}{11} \right)^2$$

$$= 3598.81 - 3589.09$$

$$= 9.72 \text{ inches}$$

$$S = \sqrt{\text{variance}}$$

$$S = \sqrt{9.72}$$

$$S = 3.11 \text{ inches}$$

13.(i) A hospital is six storeyed building. The number of rooms in each storey is 35, 32, 31, 34, 38, 33.

Find the standard deviation in the data

SOL:

x	x ²
35	1225
32	1024
31	961
34	1156
38	1444
33	1089
$\sum x = 203$	$\sum x^2 = 6899$

$$\begin{aligned}
 S &= \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2} \\
 &= \sqrt{\frac{6899}{6} - \left(\frac{203}{6}\right)^2} \\
 &= \sqrt{1149.83 - 1144.69}
 \end{aligned}$$

$$= \sqrt{5.14}$$

$$= 2.26$$

$$S = 2.26$$

ii. Subtract 5 from each number in the above data, then find standard deviation in new distribution thus obtained

SOL: Let $y = x - 5$

y	y ²
30	900
27	729
26	676
29	841
33	1089
28	784
$\sum y = 173$	$\sum y^2 = 5019$

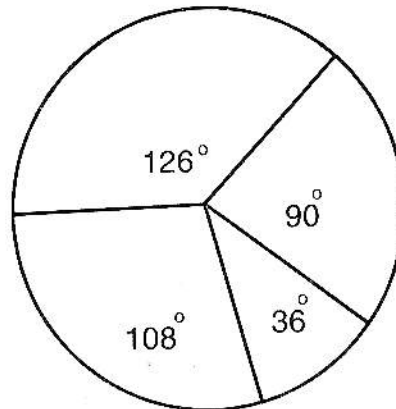
$$\begin{aligned}
 S &= \sqrt{\frac{\sum y^2}{n} - \left(\frac{\sum y}{n}\right)^2} \\
 &= \sqrt{\frac{5019}{6} - \left(\frac{173}{6}\right)^2} \\
 &= \sqrt{836.5 - 831.36} \\
 &= \sqrt{5.14} \\
 &= 2.26 \\
 S &= 2.26
 \end{aligned}$$

Q.14 Abdullah spends 35% of his income on food, 25% on electricity and clothing, 10% on his friends and the rest on miscellaneous items. Represent the income by a pie - diagram.

SOL: We want to find a pie-diagram. First we will find a table

ITEM	EXPENDITURE	ANGLE OF SECTOR
Food	35	$\frac{35}{100} \times 360 = 126^\circ$
Electricity and Clothing	25	$\frac{25}{100} \times 360 = 90^\circ$
Friends	10	$\frac{10}{100} \times 360 = 36^\circ$
Miscellaneous	30	$\frac{30}{100} \times 360 = 108^\circ$
	100	360°

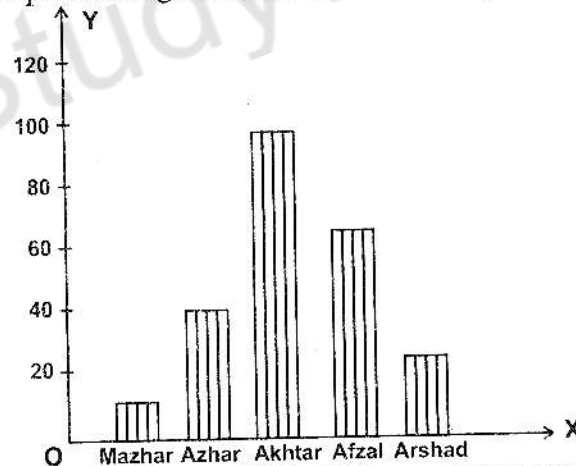
The req. pie-graph is



Q.15 Following are the runs made by five students at a sports day show the runs among the students by a simple bar chart.

Students	Mazhar	Azhar	Akhtar	Afzal	Arshad
Runs	12	40	95	65	24

SOL: The required simple bar diagram of the sums scored by 5 students at sports day is as:



Q.16 The two partners Mr. Aslam and Ms. Kalsoom run a company, in the following data the weekly wages in (Rs.) of employees who work in the company are given:

Wages (Rs.)	600-700	700-800	800-900	900-1000	1000-1100	1100-1200	1200-1300
f	3	5	7	11	21	6	2

Calculate (i) Arithmetic mean if $D = X - 950$

Wages	f	x	$D = x - 950$	fD	$u = \frac{x-900}{100}$	fu
600-700	3	650	-300	-900	-2.5	-7.5
700-800	5	750	-200	-1000	-1.5	-7.5
800-900	7	850	-100	-700	-0.5	-3.5
900-1000	11	950	0	0	0.5	5.5
1000-1100	21	1050	100	2100	1.5	31.5

1100-1200	6	1150	200	1200	2.5	15
1200-1300	2	1250	300	600	3.5	7
Σ	55			1300		40.50

$$\begin{aligned} \text{i. } \bar{x} &= A + \frac{\sum fd}{\sum f} \\ &= 950 + \frac{1300}{55} \\ &= 950 + 23.64 \\ \bar{x} &= 973.64 \end{aligned}$$

$$\begin{aligned} \text{ii. } \bar{x} &= A + \frac{\sum fu}{\sum f} \times h \\ &= 900 + \frac{40.50}{55} \times 100 \\ &= 900 + 73.64 \\ &= 973.64 \end{aligned}$$

EXAMPLES

Example 1: Following are the numbers of telephone calls made in a week to 30 teachers of high school.

5, 8, 11, 25, 13, 16, 20, 17, 15, 16, 30, 21, 14, 18, 19, 6, 22, 26, 15, 19, 35, 29, 31, 23, 25, 20, 10, 9, 7, 26

Using 5 as class interval, make a frequency distribution by Tally method.

Solution: Largest observation = 35 and smallest observation = 5

Class Interval	Tally Marks	Frequency (f)
5 - 9		5
10 - 14		4
15 - 19		8
20 - 24		5
25 - 29		5
30 - 34		2
35 - 39	I	1
Total	---	30

Example 2: Following are the ages (in months) of 37 students of a school.

125, 126, 131, 129, 130, 128, 123, 124, 136, 133, 141, 138, 148, 151, 140, 144, 145, 144, 150, 150, 143, 146, 156, 150, 152, 149, 155, 164, 170, 160, 161, 170, 160, 161, 159, 157, 154, 166

Put the above information in 10 groups or classes by direct method.

Solution: Highest Value = 170; Lowest Value = 123

Difference = Range = 170 - 123 = 47

Size of the Class Interval $\frac{47}{10(\text{groups})} = 4.7$ or 5 approx

Class Interval	Ages (in month)	Frequency
123 - 127	123, 126, 124, 125	4
128 - 132	131, 129, 128, 130, 128	5
133 - 137	133, 136	2
138 - 142	138, 140, 141	

143 - 147	143, 144, 145, 146, 146	6
148 - 152	148, 150, 151, 149, 150, 150, 152	7
153 - 157	155, 151, 157, 154	4
158 - 163	159, 160, 161	3
163 - 167	164, 166	2
168 - 172	170	1
Total	----	37

Example 3: Following are the number of children born to 40 women up to the age of 30 year. Make a discrete frequency distribution.
2, 2, 0, 1, 3, 4, 3, 5, 4, 2, 3, 1, 4, 5, 1, 1, 0, 2, 4, 5, 3, 5, 2, 3, 4, 3, 5, 3, 4, 4, 5, 3, 2, 5, 3, 4, 3, 5, 3, 4,

Solution: As figure is discrete variable so, we make a discrete frequency distribution

Children	Numbers	Frequency
0	0, 0	2
1	1, 1, 1, 1	4
2	2, 2, 2, 2, 2, 2	6
3	3, 3, 3, 3, 3, 3, 3, 3, 3, 3,	11
4	4, 4, 4, 4, 4, 4, 4, 4, 4	9
5	5, 5, 5, 5, 5, 5, 5, 5	8
Total	(Entries)	40

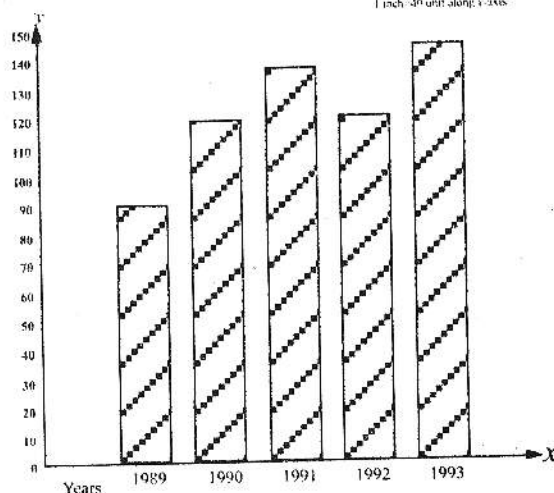
Example 4: The following data represent the number of students who passed matriculation examination in first division of a high school during the years 1989 to 1993. Represent the data by a simple bar diagram.

Years	1989	1990	1991	1992	1993
Student	90	120	138	121	145

Solution:

Title: Graph showing Simple Bar diagram for the number of students who passed matriculation examination in first division during the years 1989 - 1993

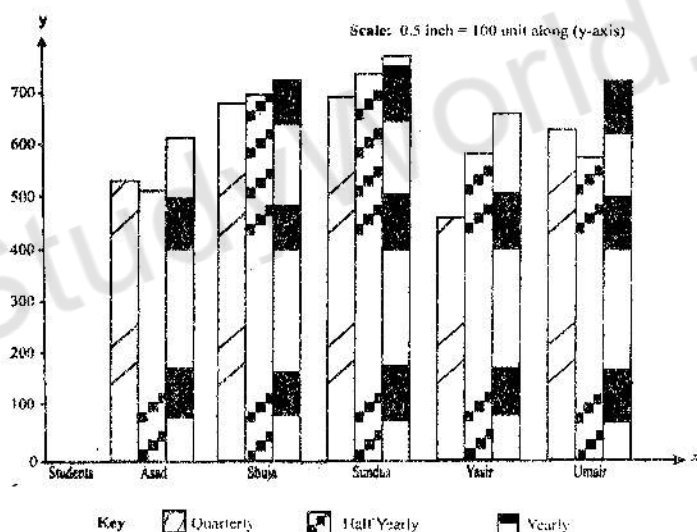
Scale:
1 inch = 40 unit along y-axis



Example 5: The following are the marks obtained out of 800 by five students in quarterly, half yearly and yearly examinations. Represent the figures by a multiple bar diagram

Name of Student	Quarterly	Half Yearly	Yearly
Asad	525	500	600
Shuja	675	690	710
Sundas	680	725	760
Yasir	455	570	650
Umair	620	566	705

Solution: **Title:** Graph showing Multiple Bar Diagram Comparison of marks obtained by five students.



Example 6: The table given below shows the population of six classes of a high school. Represent the data by a Pie-diagram.

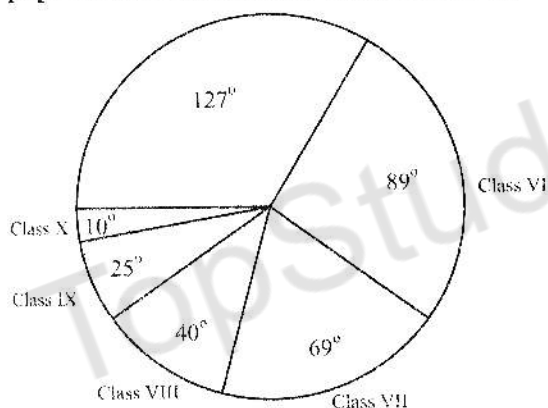
Class	Population
V	300
VI	210
VII	162
VIII	95
IX	60
X	23
Total	850

Solution:

Class	Population	Angle of the sector
V	300	$\frac{300}{850} \times 360^\circ = 127^\circ$ approx.
VI	210	$\frac{210}{850} \times 360^\circ = 88.94^\circ = 89^\circ$ approx.

VII	162	$\frac{162}{850} \times 360^\circ = 68.61^\circ = 69^\circ$ approx.
VIII	95	$\frac{95}{850} \times 360^\circ = 40.23^\circ = 40^\circ$ approx.
IX	60	$\frac{60}{850} \times 360^\circ = 24.41^\circ = 25^\circ$ approx.
X	23	$\frac{23}{850} \times 360^\circ = 9.74^\circ = 10^\circ$ approx.
Total	850	360°

Title: Pie-diagram showing the population of six classes of high school.



Example 7: The grades of a student in five examinations were 64, 75, 81, 87, 90. Find the arithmetic mean of the grades.

Solution: A.M. = $\bar{x} = \frac{\sum x}{n}$

or $= \frac{64 + 75 + 81 + 87 + 90}{5}$

or $\bar{x} = \frac{397}{5} = 79.40$

Example 8: Given below are the marks out of 100 obtained by 100 students in examination. Find the arithmetic mean.

Marks	Frequency
30 - 35	14
35 - 40	16
40 - 45	18
45 - 50	23
50 - 55	18
55 - 60	11
Total	100

Solution:

Marks	Midpoint (x)	Frequency (y)	fx
30 - 35	32.5	14	445.0
35 - 40	37.5	16	600.0
40 - 45	42.5	18	765.0
45 - 50	47.5	23	1092.0
50 - 55	52.5	18	945.0
55 - 60	57.5	11	632.0
Total	---	100 = $\sum f$	4490 = $\sum fx$

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$= \frac{4490}{100}$$

or $\bar{x} = 44.90$

Example 9: Find the arithmetic mean using short formula for the runs made by a Batsman Sharique.

Runs: 40, 45, 50, 52, 50, 60, 56, 70

Solution: Taking deviations from $A = 52$

X	40	45	50	52	50	60	56	70
$D = x - A$	-12	-7	-2	0	-2	8	4	18

Now: $\Sigma D = -23 + 30 = 7$

$$\therefore \bar{x} = A + \frac{\Sigma D}{n}$$

$$\therefore \bar{x} = 52 + \frac{7}{8} = 52 + 0.875 = 52.88 \text{ or } 53 \text{ runs.}$$

Example 10: Deviations from 12.5 of ten different values are 6, -2, 3, 5, 9, 8, 7, -5.5, 14, 11.3, -6.8, -4.2, find the arithmetic mean.

Solution: Deviations from 12.5 are:

6, -2, 3, 5, 9, 8, 7, -5.5, 14, 11.3, -6.8, -4.2

Now $\Sigma D = -18.5 + 52.5 = 34$ Also $A = 12.5$, using the formula

We have,

$$\bar{X} = A + \frac{\Sigma D}{n} = 12.5 + \frac{34}{10}$$

$$\text{or } \bar{X} = 12.5 + 3.4 = 15.9$$

Example 11: The heights (in inches) of 200 school boys are recorded in the following frequency distribution. Find the mean height by short formula.

Height (x) (in Inches)	51	52	53	54	55	56	57	58	59	60
Frequency (f)	2	5	8	24	55	45	38	16	6	1

Solution: F

Heights (x) in inches	Frequency (f)	$A = 55$ $D = x - A$	fD
51	2	-4	-8
52	5	-3	-15
53	8	-2	-16
54	24	-1	-24
$A \rightarrow 55$	55	0	0
56	45	1	45
57	38	2	76
58	16	3	48
59	6	4	24
60	1	5	5
Total	$\Sigma f = 200$	$\Sigma fD = 198 - 63 = 135$	

Now using the formula (2), we get

$$\bar{X} = A + \frac{\Sigma fD}{\Sigma f}$$

$$\text{or } \bar{x} = 55 + \frac{135}{200}$$

$$\text{or } \bar{x} = 55 + 0.675$$

$$\therefore \bar{x} = 55.68 \text{ inches approx.}$$

Example 12: Five workers do their job at a big general store. Their daily wages are Rs. 300, 350, 400, 500. Calculate the mean wages by coding method.

Solution: By using the formula we have

$$\bar{x} = A + \frac{\sum U}{n} \times h$$

Here, let $A = 450$, and $h = 50$ Then

x	300	350	400	450	500	---
$U = \frac{x-A}{h}$	-3	-2	-1	0	1	$\Sigma U = -5$

$$\bar{X} = A + \frac{\sum U}{n} \times h$$

$$\text{or } \bar{X} = 450 + \frac{(-5)}{5} \times 50$$

$$\therefore \bar{X} = 450 - 50 = 400 \text{ (Rs.)}$$

Example 13: Given below is a frequency distribution of the weights in Kg of 130 students. Find the mean weight using coding method.

Weights (Kg)	40 - 44	44 - 48	48 - 52	52 - 56	56 - 60	60 - 64
Frequency	5	13	45	32	26	9

Solution:

Weight (Kg)	Frequency (f)	x	$U = \frac{x-A}{h}$	fU
40 - 44	5	42	-2	-10
44 - 48	13	46	-1	-13
48 - 52	45	50 (A)	0	0
52 - 56	32	54	+1	32
56 - 60	26	58	+2	52
60 - 64	9	62	+3	27
----	$130 = \Sigma f$	---	---	$88 = \Sigma fU$

Now using the formula (4), we get

$$\bar{X} = A + \frac{\Sigma fu}{\Sigma f} \times h$$

Taking $A = 50$ and $h = 4$, we have

$$\text{or } \bar{X} = 50 + \frac{88}{130} \times 4$$

$$\text{or } \bar{X} = 50 + 2.71 \text{ Kg}$$

$$\bar{X} = 52.71 \text{ Kg}$$

Example 14: For the following information, find the arithmetic mean.

(i) If $D = x - 10$, $\sum D = 50$ and $n = 5$

(ii) If $U = \frac{x-15}{3}$, $\sum U = 25$ and $n = 5$

(iii) If $D = x - 100$, $\sum fD = 200$, $\sum f = 50$

(iv) If $U = \frac{x-100}{3}$, $\sum fU = 240$, $\sum f = 100$

Solution:

(i) Since $D = x - A$, therefore, $A = 10$

Hence $\bar{X} = A + \frac{\sum D}{n} = 10 + \frac{50}{5} = 20$

(ii) Since $U = \frac{x-A}{h}$ therefore, $A = 15$ and $h = 3$

Hence $\bar{X} = A + \frac{\sum U}{n} \times h = 15 + \frac{25}{5} \times 3 = 15 + 15 = 30$

(iii) Since $D = x - A$, therefore, $A = 100$

Hence $\bar{X} = A + \frac{\sum fD}{\sum f} = 100 + \frac{200}{50} = 104$

(iv) Since $U = \frac{x-A}{h}$ therefore, $A = 100$ and $h = 5$

Hence $\bar{X} = A + \frac{\sum fU}{\sum f} \times h = 100 + \frac{240}{100} \times 5$

$$\therefore \bar{X} = 100 + 12 = 112$$

Example 15: Ten students each from class 5 section A and B of a well reputed school were taken randomly. Their weights were measured in Kg and recorded as given below

Weights (Kg) Section A	30	28	32	29.5	35	34	31	33	40	37.5
Weights (Kg) Section B	35	31.5	34.5	35	32.8	38	29.5	36	36.5	34

(i) Compute the mean weight for section A and B

(ii) Conclude which section is better on Average?

Solution:

(i) We find arithmetic mean for both the sections by direct method.
(Any method can be applied)

As number of observation $n = 10$			
and $\bar{X}_{(A)} = \frac{\sum X_{(A)}}{n}$		$X_{(A)}$	$X_{(B)}$
$\therefore \bar{X}_{(A)} = \frac{330}{10} = 33\text{Kg}$		30	35
and $\bar{X}_{(B)} = \frac{\sum X_{(B)}}{n}$		28	31.5
$\therefore \bar{X}_{(B)} = \frac{342.8}{10} = 34.28\text{Kg}$		32	34.5
		29.5	35
		35	32.8
		34	38
		31	29.5
		33	36
		40	36.5
		37.5	34
		$330 = \sum X_{(A)}$	$342.8 = \sum X_{(B)}$

(ii) We have seen from the results that $\bar{X}_{(B)}$ is greater than $\bar{X}_{(A)}$. Therefore, we conclude that section B is better on the average.

Example 16: The following are the scores made by a batsman
8, 12, 18, 16, 5, 20 Find the median of the data.

Solution: Writing the scores in ascending order, we have

5, 8, 12, 13, 16, 20

As n is odd

$$\therefore \text{Median} = \left(\frac{n+1}{2}\right)\text{th item} = \left(\frac{7+1}{2}\right) = 4\text{th item}$$

Hence, 13 is the median

Example 17: Following are the marks (out of 100) obtained by 10 students in English.

23, 15, 35, 48, 41, 5, 8, 9, 11, 51

Find the median of the data

Solution: Arranging the data in ascending order.

5, 8, 9, 11, 15, 23, 35, 41, 48, 51

As number of observation is even.

$$\therefore \text{Median} = \frac{1}{2} \left(\frac{n}{2} \text{th item} + \frac{n+2}{2} \text{th item} \right)$$

$$\text{As } \frac{n}{2} = \frac{10}{2} = 5 \text{ and } \frac{n+2}{2} = \frac{12}{2} = 6$$

$$\text{Median} = \frac{1}{2} [5\text{th item} + 6\text{th item}]$$

$$\text{or Median} = \frac{1}{2} [15 + 23] = \frac{38}{2} = 19$$

Example 18: The heights of 100 student, measured to the nearest (inches) are given in the following table. Find the median.

Heights (Inches)	62.5-63.5	63.5-64.5	64.5-65.5	65.5-66.5	66.5-67.5	67.5-68.5	68.5-69.5	69.5-70.5	70.5-71.5
No. of Students	4	6	10	20	30	13	12	3	2

Solution: In the above data, class boundaries have been given.

Heights (inches)	Frequencies (f)	C.F.
62.5-63.5	4	4

63.5-64.5	6	6 + 4 = 10
64.5-65.5	10	10 + 10 = 20
65.5-66.5	20	20 + 20 = 40 → C
66.5 - 67.5	30	30 + 40 = 70 →
67.5-68.5	13	13 + 70 = 83
68.5-69.5	12	12 + 83 = 95
69.5-70.5	3	3 + 95 = 98
70.5-71.5	2	2 + 98 = 100 → n
Total	100Σf	---

Median group

Here $n = 100$ so, $\frac{n}{2} = \frac{100}{2} = 50$

50th item lies in 66.5 - 67.5

$$\therefore \text{Median} = l + \frac{h}{f} \left(\frac{n}{2} - c \right)$$

$$= 66.5 + \frac{1}{30} (50 - 40) \quad (\text{putting the value})$$

$$= 66.5 + \frac{10}{30}$$

$$= 66.5 + 0.33$$

$$\therefore \text{Median} = 66.83 \text{ inches}$$

Example 19: Following are weights (in pounds) of 50 students. Find the median weight.

Weights (lbs)	110-114	115-119	120-124	125-129	130-134
No. of Students (f)	5	12	23	6	4

Solution: As class boundaries are not given so, first of all we make class boundaries by the usual procedure.

Weight (lbs)	Frequency	Class Boundaries	C.F
110 - 114	5	109.5 - 114.5	5
115 - 119	12	114.5 - 119.5	17 → c
120 - 124	23	119.5 - 124.5	40
125 - 129	6	124.5 - 129.5	46
130 - 134	4	129.5 - 134.5	50 → n
Total	50 = Σf	----	----

Median group

Here $n = 50$ so, $\frac{n}{2} = \frac{50}{2} = 25$, 25th item lies in 119.5 - 124.5

$$\text{Median} = l + \frac{h}{f} \left(\frac{n}{2} - c \right)$$

$$= 119.5 + \frac{5}{23} (25 - 17)$$

(Putting the values)

$$= 119.5 + \frac{40}{23}$$

$$= 119.5 + 1.74$$

\therefore Median = 121.24 lbs.

Example 20: Following are the heights in (inches) of 40 students in class 8.

Heights (inches)	Frequency (f)
48 - 50	5
50 - 52	7 $\rightarrow f_1$
52 - 54	10 $\rightarrow f_m$, here $h = 2$
54 - 56	9 $\rightarrow f_2$
56 - 58	6
58 - 60	3
Total	40 (Σf)

Solution: In the above data, class boundaries have been given. Using the formula for ground data we find mode as.

$$\text{Mode} = l + \frac{(f_m - f_1) \times h}{(f_m - f_1) + (f_m - f_2)}$$

$$\text{or Mode} = 52 + \frac{(10 - 7) \times 2}{(10 - 7) + (10 - 9)}$$

$$\text{or Mode} = 52 + \frac{3 \times 2}{3 + 1} = 52 + \frac{6}{4}$$

$$\text{or Mode} = 52 + 1.5 = 53.5 \text{ (inches)}$$

Example 21: The weights in Kg of 60 students of a class are given in the following grouped data.

Weight (kg)	40 - 42	43 - 45	46 - 48	49 - 51	52 - 54
No. of Student	10	12	30	6	2

Find mode of the data

Solution: As class boundaries are not given so, first we make class boundaries by usual procedure.

Weights (kgs)	Frequencies	Class Boundaries
40 - 42	10	39.5 - 42.5
43 - 45	12 $\rightarrow f_1$	42.5 - 45.5
46 - 48	30 $\rightarrow f_m$	45.5 - 48.5 \rightarrow
49 - 51	6 $\rightarrow f_2$	48.5 - 51.5
52 - 54	2	51.5 - 54.5
Total	60 (Σf)	-----

Mode Group

As the frequencies is 30 which lies in (45.5 - 48.5) so, the modal group is (45.5 - 48.5). Using the formula we find:

$$\text{Mode} = l + \frac{(f_m - f_1) \times h}{(f_m - f_1) + (f_m - f_2)}$$

$$\text{or Mode} = 45.5 + \frac{(30 - 12) \times 3}{(30 - 12) + (30 - 6)}$$

$$\text{or Mode} = 45.5 + \frac{54}{18 + 24} = 45.5 + \frac{54}{42}$$

$$= 45.5 + 1.29 = 46.79 \text{ Kgs}$$

Note: An approximate value of the mode can be obtained by using an empirical relation;

$$\text{Mode} = 3 \text{ median} - 2 \text{ mean}$$

Example 22: Following figures show the profit (in Rs) of then shopkeepers:

70, 100, 150, 130, 140, 150, 90, 60, 110, 600

Find range of the above data

Solution: Here $x_m = 600$ and $x_l = 60$

$$\therefore \text{Range (R)} = x_m - x_l$$

$$\therefore R = 600 - 60 = 540$$

Example 23: A set contains the following numbers:

3.5, 4, 4.5, 9, 8, 3.1, 11, 10.5, 8.5

Find is Range

Solution: Here $x_m = 11$ and $x_l = 3.1$

$$R = x_m - x_l$$

$$R = 11 - 3.1 = 7.9$$

Example 24: Ten students took a test Mathematics. They got marks (out of 100) as: 46, 50, 52, 60, 63, 64, 51, 55, 66

Solution: (i) Using the proper formulas

$$S^2 = \frac{\sum (x - \bar{x})^2}{n}$$

Here, we first find \bar{x} of the data, then take the sum of squares of the deviation from (\bar{x})

Mark (x)	$x - \bar{x}$	$(x - \bar{x})^2$
46	-10.8	116.64
50	-6.8	46.42
52	-4.8	23.40
60	3.2	10.29
63	+6.2	38.44
64	+7.2	51.84
51	-5.8	33.64
61	+4.2	17.64
55	-1.8	3.24
66	+9.2	84.64
Sum : 568	----	425.60

$$\bar{x} = \frac{\sum x}{n} = \frac{568}{10} = 56.8 \text{ marks}$$

$$\text{and } S^2 = \frac{\sum (x - \bar{x})^2}{n}$$

$$\text{or } S^2 = \frac{425.60}{10}$$

$$\text{or } S^2 = 42.56 \text{ marks}$$

(ii) Using Direct Method

Marks (x)	x^2
46	2116
50	2500
52	2704
60	3600
63	3969

64	4096
51	2601
61	3721
55	3025
66	4356
Sum : 568	32688

By direct formula, we have

$$S^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n} \right)^2$$

$$\text{or } S^2 = \frac{32688}{10} - \left(\frac{568}{10} \right)^2$$

Putting the values

$$\text{or } S^2 = 3268.80 - 3226.24$$

or $S^2 = 42.56$ marks

Example 25: Five students got prize (in Rs.) 12, 16, 10, 14, 8 in making runs at a sport day. Find the sum of squares of deviations from mean (\bar{x}). Also find standard deviation.

Solution:

$$As, \bar{x} = \frac{\sum x}{n}$$

$$\therefore \bar{x} = \frac{60}{5} = 12 \text{ and}$$

$$S = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

Mark (x)	$(x - \bar{x})$	$(x - \bar{x})^2$
12	0	0
16	4	16
10	-2	4
14	+2	4
8	-4	16
	0	40

Thus the sum of squares of deviations taken from mean (\bar{x}) is 40

$$\text{i.e. } \sum (x - \bar{x})^2 = 40$$

Keep the following criteria mind for the measurement of spread of observation that:

- $(\bar{x} \pm 1S)$ covers observations approximately 68.27% of the measurements.
- $(\bar{x} \pm 2S)$ cover observations approximately 95.45% of the measurements.
- $(\bar{x} \pm 3S)$ cover observations approximately 99.73% of the measurements.

Example 26: The weight measurements (in grams) of 12 medicines are given below:

43, 54, 45, 44, 58, 47, 50, 52, 51, 45, 48, 46
Calculate its mean (\bar{x}) and standard deviation (s).

Find percentage (%) of observations lying within $(\bar{x} \pm 1S)$,

$(\bar{x} \pm 2S)$ and $(\bar{x} \pm 3S)$ limits.

Solution:

Medicines (in grams) (x)	x^2
43	1849
54	2916
45	2025
44	1936
58	3364
47	2209
50	2500
52	2704
51	2601

45	2025
48	2304
46	2116
Sum : 583	28549

We first find \bar{x} and S

$$\bar{x} = \frac{\sum x}{n} = \frac{283}{12}$$

$$= 48.58 \text{ and}$$

$$S.D = S = \sqrt{\frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2}$$

$$= \sqrt{\frac{28549}{12} - \left(\frac{583}{12}\right)^2}$$

$$\text{or } S = \sqrt{2370.08 - 2360.34}$$

$$\text{or } S = \sqrt{18.74} = 4.33$$

Now we find the limits $(\bar{x} \pm 1S)$

$$(a) \quad (\bar{x} \pm 1S) = 48.58 \pm 4.33$$

$$= 44.25 \text{ and } 52.91$$

There are eight values within lying within the limits 44.25 and 52.91

%age of value lying within

$$(\bar{x} \pm 1S) = \frac{8}{12} \times 100$$

$$= 66.66 \text{ or } 67\%$$

(b) We find the limits $\bar{x} \pm 2S$

$$\bar{x} \pm 2S = 48.59 \pm 2(4.33)$$

$$= 39.92 \text{ and } 57.24$$

There are eleven values lying within the limits 39.92 and 57.24

%age of values lying within $(\bar{x} \pm 2S)$

$$= \frac{11}{22} \times 100$$

$$= 91.66 \text{ or } 92\%$$

(c) We find limits $\bar{x} \pm 3S$

$$\bar{x} \pm 3S = 48.58 \pm 3(4.33)$$

$$= 35.59 \text{ and } 61.57$$

All values lie within the limits 35.99 and 61.57

\therefore %age values lying within

$$\bar{x} \pm 3S = \frac{12}{12} \times 100$$

$$= 100\%$$

Objective

Q. 1. Four answers of each item are given from which only one is true. Tick the correct answer.

1. The numerical figures obtained from any field of study are known as

- (a) Information (b) data
- (c) Information Handling
- (d) Calculation.

2. A person who collects the data is known as _____.

- (a) Investigator (b) Teacher
- (c) Student (d) Inspector

3. Investigator must be _____.

- (a) Judge
- (b) Intelligent, reliable and responsible
- (c) Intelligent (d) Experienced

4. There are _____ main types of data.

- (a) 2 (b) 3
- (c) 4 (d) 5

5. The data directly collected from its source is called _____ data.

- (a) secondary (b) primary
- (c) constant (d) variable

6. The data which have been passed through some statistical treatment at least once is called _____ data. (Lahore Board 2010)

(a) secondary (b) primary

(c) variable (d) constant

7. Any fixed quantity that has a single value is known as.

- (a) data (b) variable
- (c) Constant (d) grouped data

8. Any characteristic whose values are always different from one individual to another is known as a _____.

- (a) variable (b) Constant
- (c) data (d) grouped data

9. Numerical facts which are obtained on the first hand and recorded as they stand are known as

- (a) data (b) ungrouped data
- (c) grouped data (d) standard deviation

10. When the data have gone through some statistical process is called _____.

- (a) grouped data (b) ungrouped data

(c) variable (d) primary data

11. The number of classes or groups should be between _____.

- (a) 4 and 10 (b) 5 and 15
(c) 2 and 8 (d) 4 and 15
12. The number of occurrences of items corresponding to the class interval is known as _____.
(a) Class mark (b) class boundaries
(c) class frequency (d) classification
13. If $U = \frac{x-100}{5}$, $\sum fu = 240$, $\sum f = 100$
then $\bar{x} =$ _____.
(a) 112 (b) 110
(c) 100 (d) 90
14. If $D = x - 100$, $\sum fD = 200$, $\sum f = 50$
then $\bar{x} =$ _____.
(a) 104 (b) 102
(c) 90 (d) 50
15. A value of variable which appears or occurs most often is called _____.
(a) Mean (b) Mode
(c) Median (d) variance
16. If $x = 10, 15, 15, 20, 20, 20, 25, 32$, the mode is _____.
(a) 15, 20 (b) 15
(c) 20 (d) 10, 25
17. $3\text{median} - 2\text{mean} =$ _____.
(a) Variance (b) Mode
(c) Zero (d) Standard deviation
18. Co-efficient variation = _____.
(a) $\frac{s}{\bar{x}}$ (b) $\frac{s}{\bar{x}} \times 100$
(c) $\frac{\bar{x}}{s} \times 100$ (d) $\frac{\bar{x}}{s}$
19. If $x = 3.5, 4, 4.5, 9.8, 3.1, 11, 10.5, 8.5$ then Range is _____.
(a) 7.9 (b) 8.9
(c) 7 (d) 4.5
20. In a series 2, 3, 5, 5, 4, 7, 9 the mode is _____.
(a) 5 (b) 3
(c) 5.5 (d) 9
21. In a series (50 – 59), the size of class interval is _____.
(a) 10 (b) 18
(c) 38 (d) 9
22. In a series 5, 30, 10, 47, 100, the range (R) is _____.
(a) 25 (b) 95
(c) 90 (d) 10
23. The sum of 5 observations is 125, its mean is _____.
(a) 25 (b) 50
(c) 75 (d) 15
24. A series contains Values 4, 4, 4, 4, 4, 4, its standard deviation is _____.
(a) 4 (b) 1
(c) 0 (d) 2.1
25. If standard deviation of a series is 6, then its variance is _____.
(a) 20 (b) 15
(c) 2 (d) 36
26. How many types of data are there?
(a) 2 (b) 3
(c) 4 (d) 5
27. In class interval (90 – 99), the lower class limit is _____.
(a) 90 (b) 99
(c) 98.5 (d) 99.5
28. In class interval (15 – 20), upper class limit is _____.
(a) 15 (b) 20
(c) 20.5 (d) 28.5
29. In class interval (50 – 59), mid point is _____.
(a) 54.4 (b) 50.5
(c) 50.9 (d) 54.5
30. In a series 1, 3, 5 median is _____.
(a) 1 (b) 3
(c) 5 (d) 9
31. The process of arranging the data into certain groups or classes having similar characteristics is called _____.

- (a) Data (b) Classification
(c) Frequency (d) Variance
32. How many types of graph of frequency distribution are there?
(a) three (b) four
(c) two (d) five
33. How many types of measure of central tendency are there?
(a) two (b) three
(c) four (d) five
34. To find A.M from grouped data, formula is _____.
(a) $\bar{x} = \frac{\sum fx}{\sum f}$ (b) $\bar{x} = \frac{\sum x}{f}$
(c) $\bar{x} = \sum f$ (d) $\bar{x} = \sum fx$
35. To find median from grouped data, formula is _____.
(a) $\ell + \frac{\frac{h}{f}(\frac{n}{2} - c)}{\frac{n}{2} - c}$ (b) $\frac{h}{f}(\frac{n}{2} - c)$
(c) $\ell + (\frac{n}{2} - c)$ (d) $\ell + \frac{n}{2}$
36. To find mode from grouped data, the formula is _____.
(a) $\ell + \frac{f_m - f_1}{(f_m - f_1) + (f_m - f_2)} \times h$
(b) $\ell + \frac{f_m - f_2}{f_m - f_1} \times h$
(c) $\frac{\sum fx}{\sum f}$ (d) None of these
37. How many types of dispersion are _____.
(a) 2 (b) 3
(c) 4 (d) 5
38. To find range R, the formula is _____.
(a) $R = x_m - x_1$ (b) $R = x_1 \times x_m$
(c) $R = x_m$ (d) $R = \frac{x_m}{x_1}$
39. Midpoint of any class interval is called _____.
(a) Class mark (b) frequency

- (c) class limits (d) Classification.
40. Class boundaries may also be obtained from the midpoints 'x' as _____ where h is difference between any two consecutive values of x.
(a) $x + \frac{h}{2}$ (b) $x \pm \frac{h}{2}$
(c) $x - \frac{h}{2}$ (d) $x \pm h$
41. The size, width or length of a class interval is the _____ between the upper or lower limits of any two consecutive classes.
(a) Difference (b) Average
(c) Midpoint (d) Division
42. _____ is the average value of the lower and upper class limits.
(a) Class interval (b) Class mark
(c) Class frequency (d) Size of class
43. To present the information in a manageable way so that useful conclusions can be drawn, is called _____.
(a) Information handling
(b) Information
(c) Data (d) Frequency
44. _____ is a graph of adjacent rectangles.
(a) Frequency polygon (b) Histogram
(c) Simple bar diagram
(d) Multiple bar diagram
45. The _____ is used for the comparison of values of different items by making the sectors of a circle.
(a) Pie diagram
(b) Simple bar diagram
(c) Multiple bar diagram (d) Histogram
46. Measure of central tendency commonly known as an ____.

